

CURIOS BITE: AN APP TO KNOW WHAT'S IN YOUR PLATE

A Project Report

Submitted By

SHAKSHI SHAH

190303105144

SANSKAR NAYTA

190303105128

SHIVAM PARMAR

190303105131

PRATHAM PATEL

190303105132

in Partial Fulfilment For the Award of the Degree of

BACHELOR OF TECHNOLOGY

COMPUTER SCIENCE & ENGINEERING

Under the Guidance of

Dr. DAXA VEKARIYA

Professor



PARUL UNIVERSITY

VADODARA

February - 2023



PARUL UNIVERSITY

CERTIFICATE

This is to Certify that Project - 3 -Subject code 203105450 of 8th Semester entitled “Curious Bite” of Group No. PUCSE_ 130 has been successfully completed by

- SANSKAR NAYTA- 190303105128
- SHIVAM PARMAR - 190303105131
- PRATHAM PATEL - 190303105132
- SHAKSHI SHAH - 190303105144

under my guidance in partial fulfillment of the Bachelor of Technology (B.TECH) in Computer Science and Engineering of Parul University in Academic Year 2022- 2023.

Date of Submission :-

Dr. Daxa Vekariya ,

Project Guide

Head of Department,

Dr. Amit Barve

CSE, PIET

DR. Kruti Sutaria

Project Coordinator:-

Parul University

External Examiner

Acknowledgements

We first take the privilege to thank **Dr. Amit Barve, Head of the Department**, Computer Science & Engineering PIET for his valuable support and guidance during the period of project implementation.

We wish to express our sincere and gratitude to our project guide **Dr. Daxa Vekariya, Professor**, Department of Computer Science & engineering, PIET, for the simulating discussions, in analyzing problems associated with our project work and for guiding us throughout the project. Project meetings were highly informative. We express our warm and sincere thanks for the encouragement, untiring guidance and the confidence she had shown in us. We are immensely indebted for her valuable guidance throughout the project.

We also thanks to all the staff members of CSE department for their valuable advice, for providing resources as and when required and thanks to our friends who supported us and helped us when we need.

Sanskar Nayta

Shivam Parmar

Pratham Patel

Shakshi Shah

Parul University,

Vadodara

Abstract

Computer Vision is an interdisciplinary field of study that helps computers to gain deep understanding of digital images or videos and develop techniques to identify other similar images or displays based on the understandings. It has numerous applications like image recognition, object detection and image creation. Object detection is one of the most vital research options for computer vision. This technique detects the semantics of the object that belongs to a particular class in digital images or videos. These algorithms generally leverage machine learning or deep learning to provide significant and desirable results. Deep learning has gained popularity for automatic extraction of hidden knowledge from data. With deep learning, a lot of new applications of computer vision techniques have been introduced and are now becoming parts of our everyday lives. These include face recognition and indexing, photo stylization or machine vision in self-driving cars. Object detection using deep learning has given excellent results however there are many problems with images in real-world shooting scenario as noise, rotating jitter or blurring. These issues have a great impact on object detection techniques. Deep neural networks (DNNs) have been demonstrated to achieve superior object detection performances compared to other algorithms, with TensorFlow being one of the majors in DNN based object detection techniques in terms of both speed and accuracy.

Table of Contents

Acknowledgements	iii
Abstract	iv
Table of Contents	v
List of Figures	viii
1 Introduction	1
1.1 PROBLEM STATEMENT	1
1.2 MOTIVATION	1
1.3 SCOPE OF PROJECT	2
1.4 AIM	2
1.5 OBJECTIVE	3
2 LITERATURE SURVEY	4
2.1 REAL-TIME OBJECT DETECTION WITH TENSORFLOW	4
2.2 TRAFFIC LIGHT DETECTION USING TENSORFLOW OBJECT DETECTION FRAMEWORK	5
2.3 MOBILENETS FOR FLOWER CLASSIFICATION BY TENSORFLOW	5
2.4 FOREST MONITORING SYSTEM THROUGH RFID AND TENSORFLOW OBJECT DETECTION	6

2.5 TEXT DETECTION FOR NATURAL SCENE BASED ON MOBILENET V2 AND U-NET	8
2.6 ESTIMATING FOOD CALORIES FOR MULTIPLE-DISH FOOD PHOTOS	8
2.7 FOOD IMAGE RECOGNITION USING CNN	9
2.8 COMPARISON OF CNN TOLERANCES TO INTRA CLASS VARIETY IN FOOD RECOGNITION	10
2.9 DIGITAL IMAGE PROCESSING IN REMOTE SENSING	11
2.10 A DISTANCE MEASURING METHOD USING VISUAL IMAGE PROCESSING	12
2.11 RESEARCH AND DESIGN OF SMART HOME SPEECH RECOGNITION SYSTEM BASED ON DEEP LEARNING	13
2.12 A FOOD PHOTOGRAPHY APP WITH IMAGE RECOGNITION FOR THAI FOOD	14
2.13 EFFICIENCY ENHANCEMENT OF FOOD RECOGNITION USING ARTIFICIAL NEURAL NETWORK	14
2.14 SMARTPHONE-BASED FOOD CATEGORY AND NUTRITION QUANTITY RECOGNITION IN FOOD IMAGE WITH DEEP LEARNING ALGORITHM . .	15
2.15 COLLISION DETECTION MODEL BASED ON BOUNDING AND CONTAINING BOXES	16
2.16 RESEARCH ON OBSTACLE AVOIDANCE OF SPACE BASED ON CYLINDRICAL BOUNDING BOX	16
2.17 EVALUATION OF EXTREMIST COHESION IN A DARKNET FORUM USING ERGM AND LDA ISIS	17
2.18 USING THE KERAS MODEL FOR ACCURATE AND RAPID GENDER IDENTIFICATION THROUGH DETECTION OF FACIAL FEATURES	17
2.19 EXPERIMENTAL STUDY ON IMAGE INTERPOLATION FOR CONCEALED OBJECT DETECTION	18

2.20 REAL TIME VEHICLE DETECTION, TRACKING AND COUNTING USING RASPBERRY-PI	19
3 Project Flow and Methodology	21
3.1 PROJECT FLOW	21
3.2 METHODOLOGY	22
3.2.1 DATA COLLECTION	22
3.2.2 DATA PRE-PROCESSING	24
3.2.3 DATA LABELLING	26
3.3 MATHEMATICAL MODEL	27
3.4 MOBILENET	28
4 EXPERIMENTAL STUDY	29
4.1 EXPERIMENTAL SETUP	29
4.2 PERFORMANCE MEASURES	29
5 OVERVIEW OF AN APPLICATION	32
5.1 UNIQUE FEATURES	33
5.2 LIMITATIONS	33
5.3 TECHNOLOGIES USED:	33
5.4 DATASET COLLECTION	34
5.5 IMPORTANCE OF TENSORFLOW	34
6 CONCLUSION & FUTURE WORK	38
REFERENCES	39

List of Figures

2.1	Object Detection	4
2.2	MobileNet model	6
2.3	GSM Module	7
2.4	The architecture of food calorie estimation	9
2.5	Eg on detection on School lunch dataset	9
2.6	Timeline of neutral networks and deep learning	10
2.7	Output food images including accuracies and calories	11
2.8	CBERS-2B CCD image of an urban region of Brasilia.(a)band 3;(b)color composition R3G4B2	11
2.9	Merging SPOT images.(a)Panchromatic image;(b)multispectral image R3G2B1;(c)fused image	12
2.10	Runway Detection and Localization in Aerial Images Using Deep Learning	13
2.11	Research and Design of Smart Home Speech Graph	13
2.12	Network model	18
2.13	Count of vehicles on up and downstream	19
3.1	Flow Chart	21
3.2	Wireframe of the application	22
3.3	Source of data	24
3.4	DataSet	26
3.5	DataSet	26
3.6	Bounding box	27
3.7	Accuracy (y-axis) vs. the number of required operations (x-axis) for available configurations	28
4.1	Iou Explanation	30

5.1	Login Page	35
5.2	Allergy Page	35
5.3	Home Page	35
5.4	(a) Output (b) Output	36
5.5	(a) Output (b) Output	36
5.6	Food Dataset	37
5.7	Image Trainning	37
6.1	ORIGINALITY REPORT	43

Chapter 1

Introduction

1.1 PROBLEM STATEMENT

The aim of this challenge is to train models that can look at pictures of food products and detect the individual food products present in them. Object detection is one of the most significant search directions for computer vision. To locate objects in the image, we use object localization concept to locate more than one object in realtime systems. Food tourism is strongly emerging as a niche that complements other means to learn the culture of a destination. India is emerging as a destination for culinary tourism. With delicious flavors and fragrant spices, Indian cuisine is certainly popular around the world. Nowadays mobile application devices have become very popular and comparatively powerful, everyone wants to use or embrace themselves with the benefits of convolution networks on their mobile devices.

1.2 MOTIVATION

Food tourism is strongly emerging as a gap that enhances different manner to examine the way of life of a destination. India is emerging as a destination for culinary tourism. With delicious flavors and fragrant spices, Indian cuisine is certainly popular around the world. Indian meal with our curries, spices and with various flavors has always been treated like an art. People from around the globe come to India and try our famous cuisine. A large section of these tourists are first time visitors or do not have much knowledge about the Indian food culture. Considering the above statistics, it is not possible to cater individually to such a large crowd, hence there is a need for an application which will help these tourists who are going to try different food items. Also, these tourists don't know what they are going to eat I.e., what's are the ingredients of the item and if they are allergic to a particular item, should they eat the item. So, Curious Bite can be considered a perfect companion which aims to provide information on most of the Indian food items.

Data Bridge marketplace studies analyses that the market is growing with a CAGR of 29% within the forecast duration of 2020 to 2027 and is anticipated to attain USD 10,189.62 million by way of 2027. So, at this time of digitization, Curious Bite can be a great opportunity to gain profits and establish monitoring power in the market.

Presently, India is home to ninety million fitness-aware people (HCl's) who are looking for nutritional values in the food they devour. Stats shows that 4 in 10 consumers closely monitor what they eat and 9 claim to be on a strict diet. Many travelers can attest to suffering from some level of food poisoning in India at some stage in their travels. Hence, to get some nutritional facts on the food which a person is eating, Curious Bite is the perfect application.

1.3 SCOPE OF PROJECT

The model is designed using object detection techniques so that it will detect different Indian food items and then integrated into web application. The application is limited to 40 different food items as for now. The application can detect multiple objects at the same time. Also, the app will show some additional nutritional facts about the detected or searched item. Overall, this project is to detect all the items in the plate and estimate the total calorie value of all the items.

1.4 AIM

The aim of this project is to identify every food item that is on the plate or table and give information about the particular object using a pre-trained model and integrate this model into web-based application, Curious Bite. Curious Bite can be considered a perfect companion which aims to provide information on most of the food items. Using object detection technique, this app helps to identify different food items based on the snapshot.

The model is trained using TensorFlow. This algorithm uses deep neural networks for object detection and is one of the faster object detection algorithms out there. There is a search option in the application which helps the user to find information on a specific food item. After detecting the particular food item, the application will provide the user with some information such as nutritional value, details, ingredients, recipes and facts about the respective item.

1.5 OBJECTIVE

- The application is integrated with Google Maps for the purpose of route optimization to find nearest restaurants and also, find restaurants serving particular food item.
- Users can directly view YouTube videos of the recipe of the detected food item.
- The application will also provide users with nutritional values of the food items.
- Also, the app will show some additional nutritional facts about the detected or searched item.
- Users can add list of food items that they are allergic from and if the item they searched have the respective ingredient, the app will show a red sign stating the user that he is not advised to eat that item.
- The application also has a feature of voice integration i.e., it can help users to easily pronounce the name of the detected item.

Chapter 2

LITERATURE SURVEY

2.1 REAL-TIME OBJECT DETECTION WITH TENSORFLOW

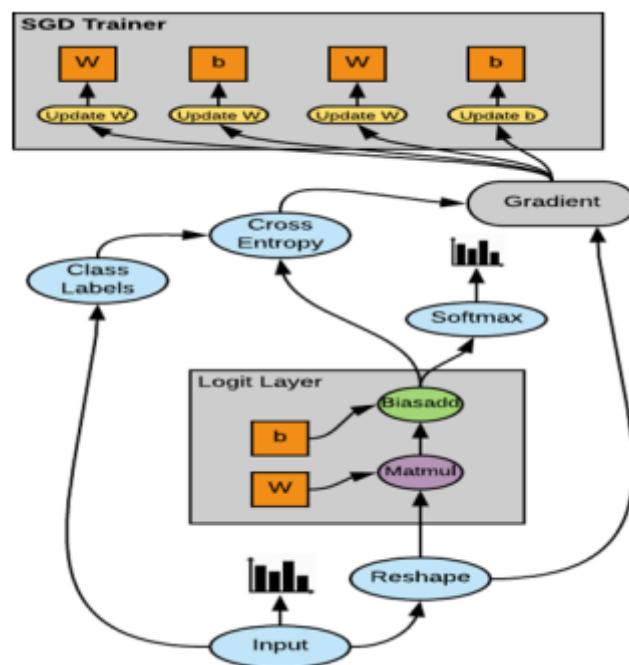


Figure 2.1: Object Detection

- The application and pervasiveness of machine learning algorithms has revolutionized the perception of computer vision problems.[2]
- These representations are represented by other simpler representations. In other words, a deep learning system can represent the concept of an image of object. Use each neuron of the equal characteristic map to extract local features at specific positions within the previous layer.[2]
- The remaining 20% is called validation data and is used to validate the data after training. After training and validating each batch, the entire data-set is tested together. It has another test

data set. This paper mainly describes how advances in CNN-based schemes have made them optimal for images. Despite the convincing results recorded so far, there is considerable room for further progress. Another type of data is tag data, which is provided at the end of the model and compared to the predicted output.[2]

2.2 TRAFFIC LIGHT DETECTION USING TENSORFLOW OBJECT DETECTION FRAMEWORK

- The fast development of automotive technology is targeted on presenting us with the high-quality features, the car's automatic driving device (ads) may be used while a person desires to switch the car into self-sufficient riding mode, or while the person can do all the paintings of driving whilst wants to transfer to computerized riding mode. I don't have faith in driving.[1]
- Machine studying is a graphical processing unit (GPU). Everyday crucial processing unit (CPU) isn't always suitable for appearing large matrix multiplication computations because the training phase can take longer. consequently, GPU-based hardware was used on this mission.[1]
- Traffic lighting fixtures were captured in numerous poses, orientations and, and perspectives. This allows deep getting to know convolutional neural networks to efficaciously learn items of hobby.[1]
- Out of a complete of 551 pics, 441 pictures were selected for schooling and a hundred and ten images have been selected for trying out.[1]
- Snap shots were cropped to square ratio and resized to 800 x 800 pixels. these images are annotated and stored in XML layout. For both fast RCNN and SSD Mobile Net V2 architectures, XML files are transformed to CSV layout. those CSV files have been converted into TF datasets which can be fed into deep mastering framework.[1]

2.3 MOBILENETS FOR FLOWER CLASSIFICATION BY TENSORFLOW

- On this paper, we proposed a stepped forward TensorFlow-based totally neural network for De-identification generation. Convolutional neural networks additionally carry out photograph adjustments with some diploma of rotation, translation, and distortion. This community avoids complicated photograph preprocessing and lets in users to directly input the original photo. accurate and fast fashions that without difficulty adapt to design wishes primarily based on cellular and embedded photo processing applications.[3]

- Section II describes existing paintings with Tensor float in constructing small models.[3]
- Section III describes essential topic, his MobileNet architecture, and the width and decision multipliers as his hyperparameters of the community. Within the above given MobileNet model, at the left aspect the labeled node “SoftMax”, is the output layer of the authentic version. MobileNets found in depth discriminating convolutional neural networks. by means of buying and selling an affordable quantity of accuracy for shrinking length and latency, we have tested a way to construct smaller and quicker MobileNets programs as flower classifiers with width and backbone multipliers.[3]
- Computing the layers at the back of the layer at once earlier than the final output layer appearing the grouping for every image takes a significant amount of time. The lower layers of the network remain unchanged, so the output can be saved and reused. Finally, we validated MobileNet’s efficiency when applied to a wide range of image datasets.[3]

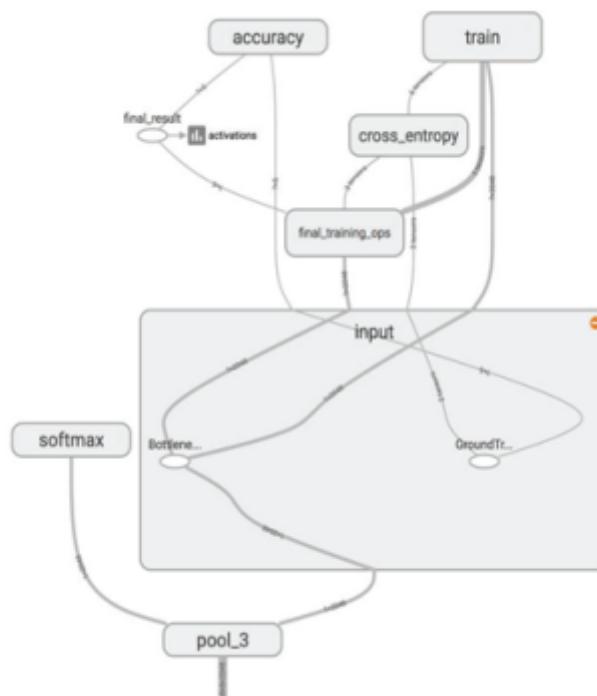


Figure 2.2: MobileNet model

2.4 FOREST MONITORING SYSTEM THROUGH RFID AND TENSORFLOW OBJECT DETECTION

- Humanity to save in this earth. one of the primary reasons for that is the invasion of trespassers into government-controlled woodland regions. And it results in unlawful activities consisting of killing wild animals and birds and cutting down trees. We spotted a few intruders coming

into the Karnataka nation Reserve Police premise from deep in the forest and poaching sandalwood timber. further, many animals are endangered and ecosystems are threatened.[4]

- Our GSM era used for verbal exchange among systems in the woods and officials. GSM is an open virtual cellular technology used for the transmission of cell voice and statistics offerings, working within the 850 MHz, 900 MHz, 1800 MHz and 1900 MHz frequency bands.[4]
- The GSM machine changed into developed as a virtual machine that makes use of Time department a couple of get right of entry to (TDMA) generation for communique purposes. Now that an authorized man or woman has been recognized, the true case of our project is complete. Then there's the fake case.[4]
- If the individual detected by way of RFID verification is incorrect, the character is recognized however without the RFID tag. The GSM module transmits photos thru any virtual media internet utility for computer and computer or in the form of cell applications. wooded area intruder.[4]
- This project will save forests and wildlife from criminals such as poachers, hunters, deforestation, and some illegal activities in forests. This is a system that can protect you from various illegal activities. Finally, I have been involved in many new technologies that have only recently been developed and have proven to be beneficial to several other works related to other technical backgrounds.[4]



Figure 2.3: GSM Module

2.5 TEXT DETECTION FOR NATURAL SCENE BASED ON MOBILENET V2 AND U-NET

- For the experimental a part of this paper, we need to categorise the pixels of the picture and finally get the masks of the textual content regions, so we use entire textual content and ICDAR 2013 as datasets inside the test. For complete text, the white area represents the textual content place and the black represents the background.[5]
- The MobileNet V2 version introduces inverse residuals and linear bottleneck structures primarily based on the MobileNet model. This not handiest speeds up the neural network model, but also reduces the complexity of the neural network. On this paper, we combine MobileNet V2 with U-internet to layout a light-weight text detector that reduces neural network complexity and drastically improves text recognition velocity.[5]
- The whole technique is especially divided into 4 elements. within the first component, the version makes use of MobileNet V2 to extract textual content features from photos. in the 2nd element, we use U-internet to carry out semantic segmentation of the picture to decide if the pixels in the photograph belong to the text place and generate a mask for the textual content region.[5]

2.6 ESTIMATING FOOD CALORIES FOR MULTIPLE-DISH FOOD PHOTOS

- Meals pictures commonly include a couple of forms of food. To understand a meals image with a couple of dishes, each dish inside the meals image desires to be recognized. meanwhile, in current years, the arrival of CNN has dramatically improved the accuracy of item detection. in this paper, we observe faster RCNN.[6]
- In this text, we recognized dishes from multiple food pics from faster R-CNN. on this test, a dataset of 21 categories of school lunch pix become created via web picture mining. It became capable of recognize food items in college lunches with excessive accuracy. similarly, we implemented this food detector to calorie estimation of ingredients from meals pics of a couple of dishes.[6]
- For evaluation that estimates meal energy from more than one meal snap shots, we also create a dataset containing images of college lunches labeled with energy. as a result of, we predicted the total energy of food from multiple dishes.[6]



Figure 2.4: The architecture of food calorie estimation



Figure 2.5: Eg on detection on School lunch dataset

2.7 FOOD IMAGE RECOGNITION USING CNN

- Technological advances and innovations such as Android phones and iPhones have also expanded the computing power of the However, people are becoming more and more conscious about their food and diet. This innovation can be used to help individuals configure different types of foods and their medicinal properties.[7]
- A new methodology is introduced for grouping food images using the convolutional neural system. The presentation of the framework is of high quality and is considered suitable from a usage point of view. In any case, CNNs require elite computing machines to examine massive multimedia datasets.[7]

- CNNs are specifically equipped to convey indirect information, increasing the amount of computation to prepare the system. In any case, the exhibition plays a big role. And once the framework is properly prepared, the framework can produce results in a short time. The image is tried on the CNN.CNN takes longer.[7]
- Because CNN is computationally expensive. We recommend a role-based approach. Misclassification occurs to overcome this hierarchical approach as the number of classes increases.[7]

2.8 COMPARISON OF CNN TOLERANCES TO INTRA CLASS VARIETY IN FOOD RECOGNITION

- Inside-class variant defines the photo variant that occurs among pics in different instructions. The similarity among samples inside the identical class is generally measured by using the intraclass correlation coefficient. A high intraclass correlation coefficient of near 1 indicates high similarity between samples in the same magnificence, and a low ICC close to 0 method the alternative.[8]
- This paper deals with the within-magnificence diversity problem for dataset Kegels Food101. Twenty-one classes with high ICC rankings had been selected. We implemented recognized convolutional neural networks consisting of ResNet, GoogleNet, MobileNet, and VGG-net with one-of-a-kind train and check possibilities to examine magnificence detection quotes.[8]

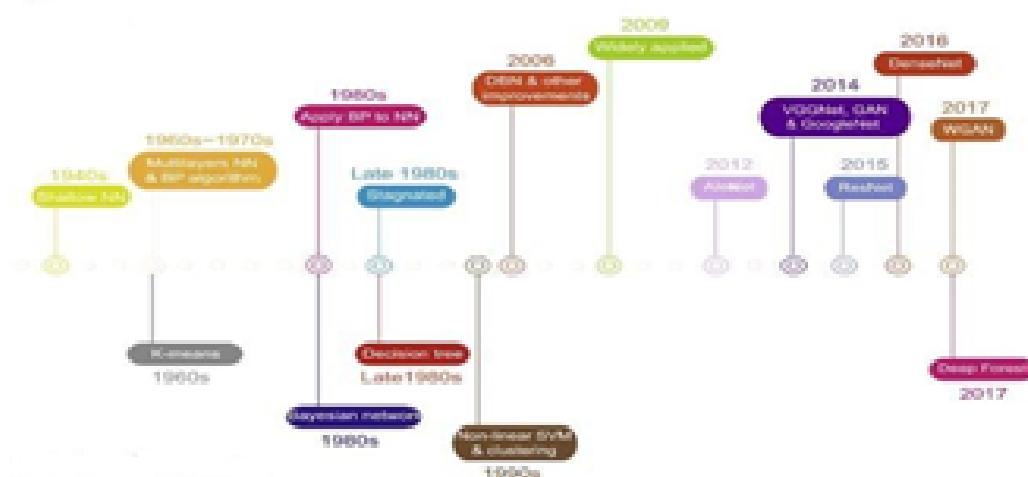


Figure 2.6: Timeline of neutral networks and deep learning

- With an increasing number of overweight people around the world, it is important to have a device that counts the range of calories in meals consumed in eating places. on this look at, we in comparison the fulfillment of deep studying algorithms across classes which might be specially difficult to define in literature meals databases.[8]



Figure 2.7: Output food images including accuracies and calories

2.9 DIGITAL IMAGE PROCESSING IN REMOTE SENSING

- Imaging systems, particularly on-board satellites, have been used in many far-off sensing applications such as city boom, deforestation, crop tracking, weather forecasting, land use mapping, and land cover mapping. provide a view. each software calls for organising a specific technique for extracting statistics from photograph data.[9]
- This white paper furnished an overview of the overall steps used to clear up remote sensing programs using picture processing techniques. A case have a look at primarily based on a city utility is supplied to demonstrate faraway sensing packages.[9]
- Attention has been directed to fundamental photo processing techniques, with the hope that the facts furnished right here will permit the interpreter or scientist to behavior research on faraway sensing packages.[9]

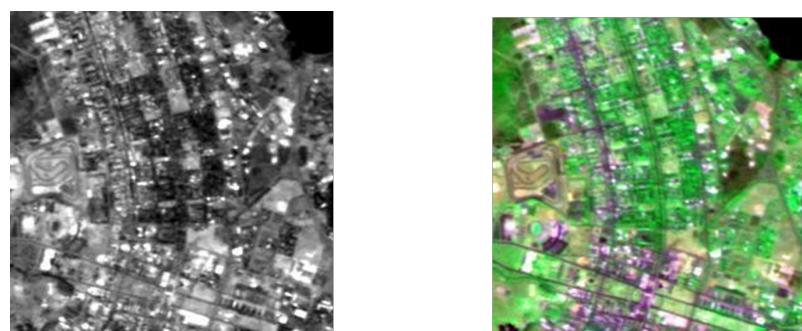


Figure 2.8: CBERS-2B CCD image of an urban region of Brasilia.(a)band 3;(b)color composition R3G4B2

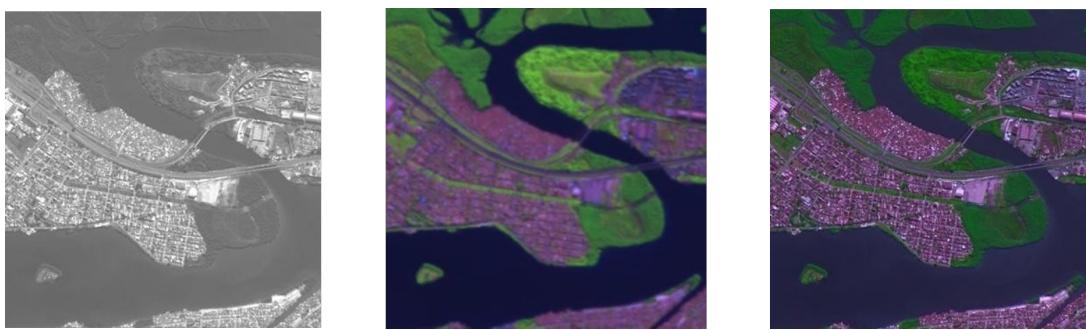


Figure 2.9: Merging SPOT images.(a)Panchromatic image;(b)multispectral image R3G2B1;(c)fused image

2.10 A DISTANCE MEASURING METHOD USING VISUAL IMAGE PROCESSING

- This paper presents a visible method to distance dimension. First, snap shots are taken at different positions. The scaling parameters among these pix can then be calculated using the Fourier-Mellin transform. The Fourier-Mellin transform is effectively used to discover the scaling and rotation parameter between pics, but the remodel itself is translation impartial.[10]
- This paper specially introduces a brand-new method to degree the distance of an item to the camera via taking benefit of the residences of the camera photo version and image processing strategies. all through processing, the Fourier-Mellin transform principle and the pinhole imaging version are used.[10]
- Experimental results show that this method shows better size results. This approach does no longer paintings nicely whilst there is an item with masses of depth facts at the optical axis of the camera because it is measured with one digicam.[10]
- I also located that there are a few restrictive conditions that limit its applicability. in the future, we plan to make suitable enhancements to the to enhance its accuracy and robustness.[10]
- Runway Detection and Localization in Aerial Imagery the usage of Deep mastering touchdown is the maximum hard phase of flight for any flight platform. the lack of an green machine has ended in numerous touchdown injuries and damage to the on-board hardware. vision-based totally systems offer a price-powerful answer for land website detection by way of offering wealthy text facts.[10]
- To this quit, this observe specializes in the accurate detection and localization of runways in aerial imagery containing disordered terrain. This facilitates air platforms, mainly his unmanned aerial vehicle (commonly called a drone), stumble on touchdown objectives (such as runways) to useful resource computerized touchdown.[10]

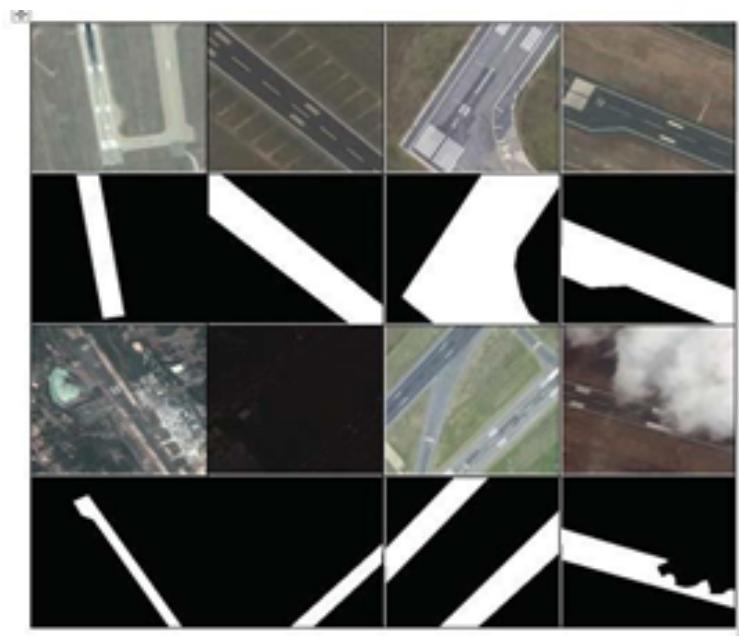


Figure 2.10: Runway Detection and Localization in Aerial Images Using Deep Learning

2.11 RESEARCH AND DESIGN OF SMART HOME SPEECH RECOGNITION SYSTEM BASED ON DEEP LEARNING

- When with the development of net data era, as society enters the technology of facts, people can now not be separated from work, look at, and lifestyles in existence. nowadays, all sorts of gadgets are directed to connection and intelligence, and the operation mode has also been modified from a touch button to his audio command control.[11]

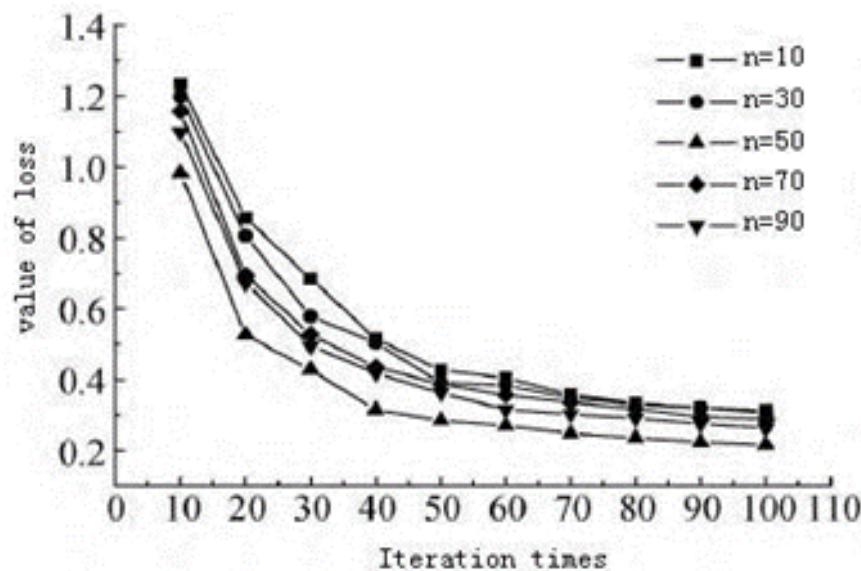


Figure 2.11: Research and Design of Smart Home Speech Graph

- From the test effects, you could see that this method shows better dimension results. This method does not paintings nicely if there is an item with lot intensity at the optical axis of the

digicam to degree with one camera. It additionally indicates that there are a few restrictions of situations that restrict the software.[11]

- Within the future, we plan to make suitable upgrades to improve accuracy and robustness. the popularity of the runway and the location laid out in air pix the usage of deep getting to know are the most difficult flight ranges for any flight platform.[11]
- With the rapid development of socio-economic system and the steady development of cutting-edge technology and generation, artificial intelligence and system studying have grow to be a research awareness and are widely used in internet seek and speech popularity.[11]
- Consequently, in the future socio-economic improvement, professionals and scholars on this discipline will deeply observe synthetic intelligence era, make complete use of its position in speech reputation machine, replace and optimize the speech popularity gadget in order that human It need to be useful in lifestyles. and manufacturing activities.[11]

2.12 A FOOD PHOTOGRAPHY APP WITH IMAGE RECOGNITION FOR THAI FOOD

- This article introduces a food photo app for smartphones that can recognize 13 types of Thai food from photos. With this feature, the application helps users calculate calories and make suggestions simply by taking a picture of the food they ate.[12]
- The application uses React Native for the frontend and Python Flask for the backend. Design a deep convolutional neural network that learns from a dataset for image recognition.[12]
- This gives us a test accuracy of about 82top 3 and top 5 ratings, the model reports a top 3 failure rate of 2.6failure rate of 1.3

2.13 EFFICIENCY ENHANCEMENT OF FOOD RECOGNITION USING ARTIFICIAL NEURAL NETWORK

- In this article, we tend to apply artificial neural networks (ANNs) to detective work tasks and food image recognition. Image recognition of foods is usually very difficult as descriptions of the wide variety of food styles. In any case, deep learning is said to be a highly intensive image recognition system in his late, and ANNs may be a dynamic way of processing deep learning.[13]
- We tend to associate his ANN with food localization and parametric perception concerns. The ANN showed fundamentally higher accuracy than the traditional support vector machine-based

routes determined manually. Additionally, I tend to note that the convolution bit indicates that the shading dictates the component extraction strategy.[13]

- For food image recognition, ANN also showed much higher accuracy than conventional methods. Radically more accurate than his traditional strategy.[13]
- In this work, we tend to present an alternative to the expected THEN, in which have been photo-confirmed removed from the pre-built THEN where the zone unit used the ILSVRC-1000 class dataset occurred in testing and achieved the highest ordering accuracy of 72.26to support grouping runs in combination with the standard option. Future work will implement the planned framework on mobile devices. To do this, we need to reduce the number of pre-trained THEN parameters containing 60 million values.[13]

2.14 SMARTPHONE-BASED FOOD CATEGORY AND NUTRITION QUANTITY RECOGNITION IN FOOD IMAGE WITH DEEP LEARNING ALGORITHM

- In keeping with similar dietary homes, foods may be divided into 6 organizations (vegetables, fruits, dairy merchandise, oils, grains and protein ingredients) to nourish the human body. however, human beings have been usually unable to apprehend the nutrients in the meals they obtained. consequently, on this paper[14]
- We suggest a deep getting to know-based totally device for education. whilst a person takes a picture of a meal with their phone, the machine acknowledges the group kind and the quantity of nutrients they devour.[14]
- Our system allows customers to be aware of the vitamins of their eating regimen and effectively manipulate their fitness the primary cause of this paper is to robotically identify vitamins in food pix. [14]
- After taking an image of the meals, the system will display the quantity of vitamins in the food to the user, so that the consumer can recognize what kind of food and what kind of they ate. in this paper, the CNN community structure to become aware of the content material of every meals from the six groups.[14]
- Compare the structure referenced on this record with other current technologies. The detection charge of CNN network structure is higher, which presents users with more accurate outcomes. within the future, we are able to acquire greater meals photos to boom the amount of schooling

and trying out statistics, optimize the proposed technique and technique, and reduce the time complexity to enhance the detection rate and the performance of mitigate.[14]

2.15 COLLISION DETECTION MODEL BASED ON BOUNDING AND CONTAINING BOXES

- We present a hierarchical model for collision detection of convex polyhedral. This model is based on the dual representation. The exterior consists of a minimal outer box that houses the polyhedron. They are Axis-Aligned Bounding Boxes (AABB) and the interior consists of up to interior boxes contained in the polyhedron.[15]
- Inner box allows faster overlap detection test, outer box allows faster non-collision test
The proposed bi-representation model was structured as a binary tree, the polyhedron was subdivided times and the original polyhedron was reconstructed through all nodes at the same level of the tree. Experimental results show that the inner box improves the collision detection process.[15]
- The result is improved collision detection efficiency. The collision detection algorithm was implemented in and validated on a set of convex polyhedral.[15]
- Intersections between contained boxes have a large impact on collision detection. The main advantage is to minimize the use of comparisons between polyhedral faces. Geometric complexity bottlenecks collision detection.[15]

2.16 RESEARCH ON OBSTACLE AVOIDANCE OF SPACE BASED ON CYLINDRICAL BOUNDING BOX

- Considering the impediment avoidance trouble of area manipulator system, a brand-new obstacle avoidance technique based totally on cylindrical bounding container version changed into proposed in step with the working environment of GIS tank detection in substation.[16]
- First, a piecewise description of the right manipulator trajectory extraction, a spatial problem to multi objective optimization that can't be determined as unknown magnitudes and undetermined parameters, a manipulator movement hassle The needs and limitations of manipulator manage in the utility and established order of the corresponding model primarily based at the analysis of Cartesian space, optimization function description is fixed in Cartesian area. Manipulator by using weighting issue.[16]
- A special environment for the mechanical arm is designed and has a bounding container to locate the cylindrical cylinder for collision detection. After the armband is first analyzed in

keeping with the set of rules for determining the preliminary role and the obstacle role of the manipulator to decide the trajectory of the manipulator corresponding to the trajectory of the manipulator, respectively, the inverse kinematics of the manipulator.[16]

- Finish To simulate mechanical arm motion in MATLAB by shifting joint angle values. After analyzing various adjustments in joint angle values on this paper, simulated the generated record with a UR simulator through the manipulator manipulate machine, and in fact read the file consistent with the trajectory of the manipulator and detected boundaries avoid.[16]

2.17 EVALUATION OF EXTREMIST COHESION IN A DARKNET FORUM USING ERGM AND LDA ISIS

- Similar extremist communities use dark web forums to network and spread news and propaganda. This article attempts to understand their network on the online forum using descriptive statistics, exponential random graph models (ERGM), and topic modeling.[17]
- Our analysis shows how bonds between active members form and grow over time. Top forum visitors were found to have higher centrality measures and her other attributes of influencers.[17]
- Use ERGM to better understand the cohesion among extremist members of the darknet forums and identify long-term trends related to overall cohesion and cohesion on issues. Using the results of ERGM modeling, we were able to compute the sum of the probabilities that two members access the most popular thread a given number of times.[17]
- While this result is not entirely conclusive in identifying top influencers, it does provide a starting point for analyzing complex and noisy networks. Understanding key influencers in your network is only pieces of the puzzle. Enriching the node with additional information is a complicated process, as many members are masked on the dark web.[17]

2.18 USING THE KERAS MODEL FOR ACCURATE AND RAPID GENDER IDENTIFICATION THROUGH DETECTION OF FACIAL FEATURES

- Using the Keras model for accurate and rapid gender identification through facial feature detection was analyzed using text or audiovisual sources. Too many people have proposed various approaches to automatically identify gender using features acquired from human bodies and behaviors.[18] The process of automatic gender determination has received a lot of praise lately due to moral and security concerns.[18]

Network model, threshold=4

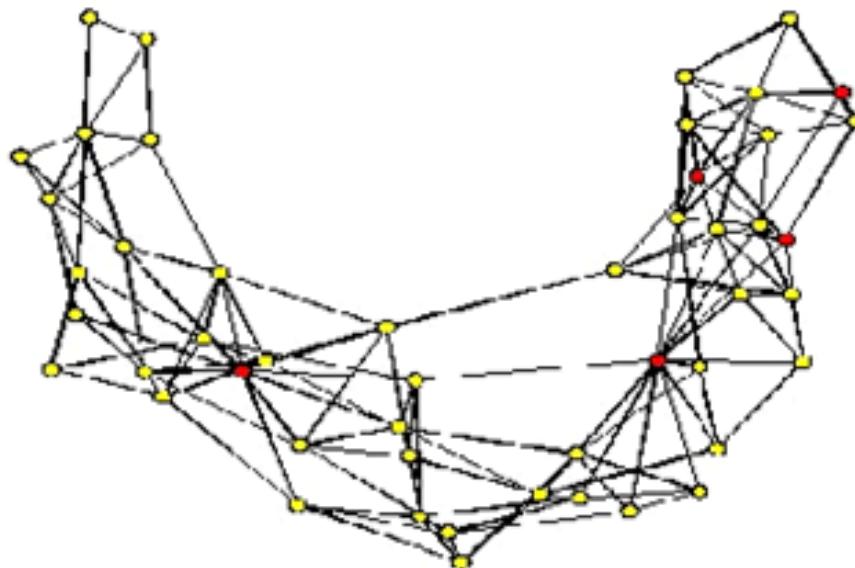


Figure 2.12: Network model

- Over the past decade, the number of photos uploaded to the Internet has grown almost exponentially. This newly discovered detail makes it easier for engineers to solve previously inaccessible or difficult-to-solve computer vision problems.[18]
- This improved the overall accuracy. This method has achieved high to moderate accuracy, depending on the properties and behavior and the technique used for precise identification. [18]
- Input can be obtained from a video camera or an image can be given as input. Here the input image is taken from the webcam. Gender is detected along with prediction accuracy. This technique is a non-invasive approach. Keras is an opensource neural network, so depending on your application, you can train this model from scratch. This greatly increases the efficiency of the system.[18]

2.19 EXPERIMENTAL STUDY ON IMAGE INTERPOLATION FOR CONCEALED OBJECT DETECTION

- Millimeter wave (MMW) imaging is extensively used in safety and military packages because of its capability to penetrate fabric and apparel [1-3]. additionally, passive MMW imaging structures can produce interpretable photos even in bad visibility conditions including fog and smoke.[19]

- A passive MMW imaging gadget acts as a distance sensor the usage of natural radiation. In assessment, energetic imaging techniques require his MMW light source to generate and emit radiation for the detection space Weighted basically Non-Oscillatory (WENO) interpolation technique.[19]
- This approach is suggested for numerical solution of partial differential equations whilst high accuracy is needed for persistent bumps or discontinuities. Non-oscillatory and weighted ENO are finite distinction or finite quantity schemes for the numerical answer of partial differential equations.[19]
- Weapon photo findings display that the 4 interpolation techniques decorate numerous features of weapon items by way of decreasing body and heritage results. performance may be evaluated in phrases of segmentation and reputation following the interpolation section. Quantitative evaluation of overall performance is left for destiny studies[19]
- We taken into consideration four interpolation methods to enhance the photo nice of hidden object detection. One is bicubic with Gaussian smooth filtering. Others are the Lagrangian method, the cubic spline method, and the WENO method. different spline and model methods may be considered for future research.[20]

2.20 REAL TIME VEHICLE DETECTION, TRACKING AND COUNTING USING RASPBERRY-PI

- It has been observed that the principal boundaries on roads are because of heavy site visitors all through rush hours, especially whilst people shuttle. a complete of 4,444 motors or gadgets have prompted avenue closures for emergency motors together with hearth trucks and rescue motors, exceeding that amount.[20]

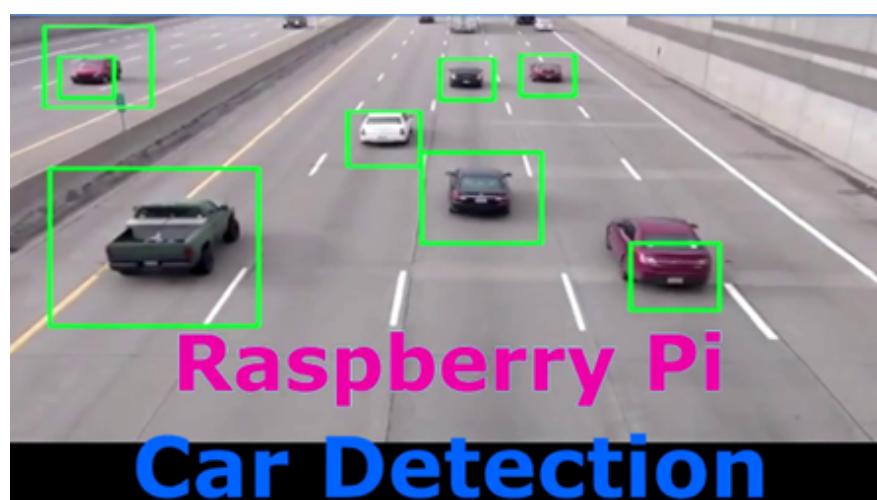


Figure 2.13: Count of vehicles on up and downstream

- Additionally wasted gasoline four,444 contributes to the country's economic increase and unsustainable environmental pollutants. The main research cognizance is on object detection and monitoring, in the end counting the quantity of motors in every surveillance place.[20]
- The want for a site visitors tracking system is to enable civil engineers and different people to plan economically and make the proper selections primarily based on car density and data derived web page from less expensive digital devices. is taking place. in addition, it offers solutions to key issues together with vehicle injuries, vehicle robbery detection, parking control, and other safety threats.[20]

Chapter 3

Project Flow and Methodology

3.1 PROJECT FLOW

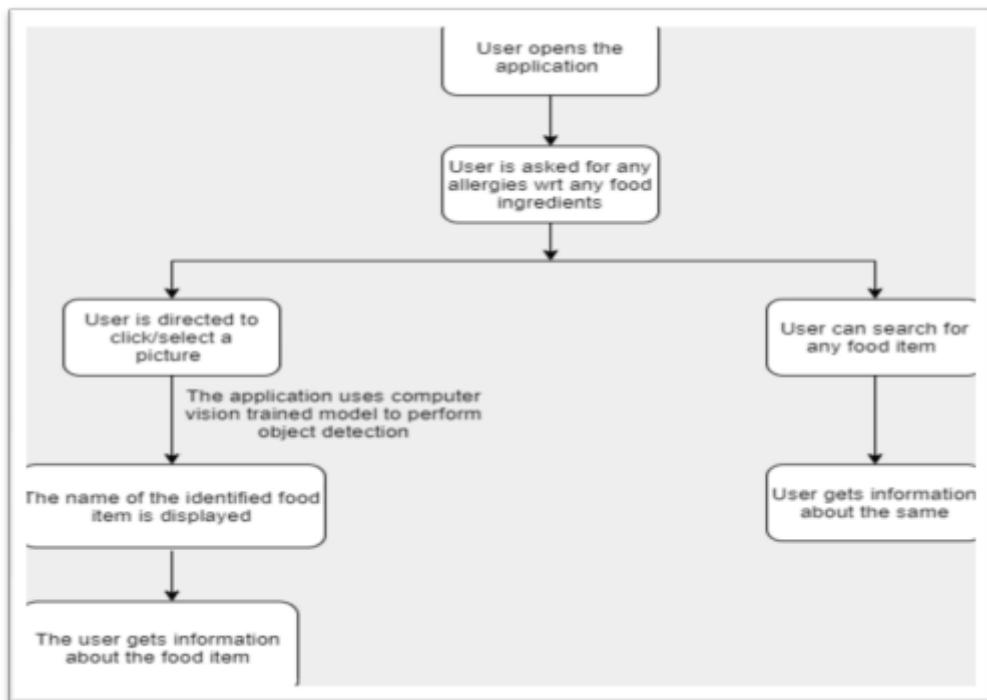


Figure 3.1: Flow Chart

- The user interface will be a web application wherein on opening the application, it will show the user a dialog box asking for name of food items they are allergic from.
- On the first screen there would be a button to allow user to select the picture he wants to scan either from camera or gallery.
- Next, user is asked to crop the image and send only the part of the image that is to be detected. Clicking on the Scan button invokes the trained TensorFlow model and will detect the inputimage.

- The name of the detected Indian Food will be displayed on the nextscreen.
- On clicking the Go button the user is then directed to a new page in which the user will find information related to the detected item with name, ingredients, YouTube link for the recipe, nutritional value chart and some facts about the respective food item.
- The application is also integrated with Google Maps and hence if the user wishes to visit any particular restaurant or nearby restaurants, they can just click on the nearby restaurant button and they will be directed to the Google Maps.

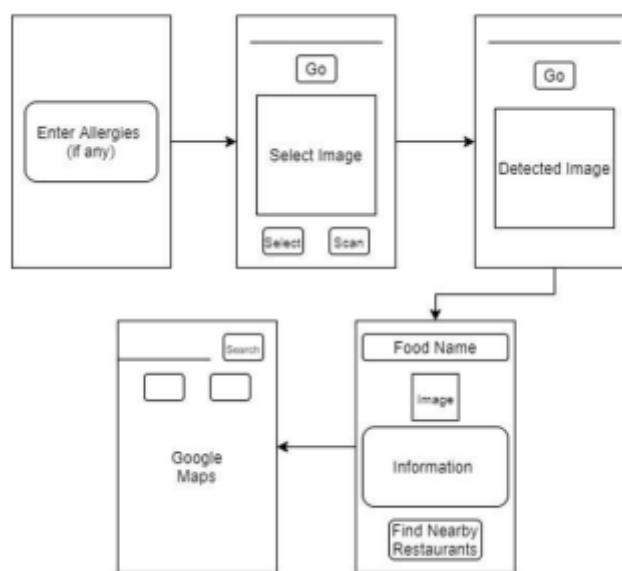


Figure 3.2: Wireframe of the application

3.2 METHODOLOGY

3.2.1 DATA COLLECTION

Data is considered the oil of industry. This is the most important part of any of the project. Most work is considered complete when data collection is performed accurately and accurately. This project required collecting images of various food items and information about each item. Various image search engines such as Google, Pinterest, and Instagram were used in the data collection process and a total of 15149 images of 40 food items were collected.

- **Web Scraping:** The basic option is to do an image search on the web and download manually. This method is not efficient as it requires a large amount of data. To solve this problem, I downloaded the Google Image Downloader extension and attached it to Chrome to download all available images at once. Image Downloader is not for sale, always free and open source, no ads or tracking algorithms. Clicking the Download button will save all selected images

to Chrome's default download directory , or a directory within it if you specified a subfolder name. However, since Google has a huge database and not all images belonged to the item we were looking for, there were many misclassified and duplicate images.

- **ImageNet:** ImageNet6 (currently only nouns), where each node of the hierarchy is represented by hundreds and thousands of images. The ImageNet data set consists of a total of about 14 million images of 21,841, different object categories. ImageNet's popular object categories include animals (fish, birds, mammals, invertebrates), plants (trees, flowers, vegetables), and activities (sports). Data is available free of charge to researchers for non-commercial purposes. Project requires Indian food such as pub bhaji, kheer, and kulfi, so I was unable to find images of most of the items on ImageNet. However, for products such as pizza, burgers, and fries, has many images available for free download.
- **Google's Open Images:** Open Images is a dataset of approximately 9 million images annotated with image-level contains a total of 16 million bounding boxes for 600 object classes on 1.9 million images, making it the largest annotated object location in the existing dataset. These records have multiple versions, V1 through V6, with V6 being the latest and largest record. Open Images Extended is a collection of sets that complement the Open Images core data set with additional images and annotations. It has a similar problem to ImageNet in that it failed to retrieve images of his Indian native foods such as Kheer and Idli. However, this method reduces the step of annotating many images.
- **Pinterest:** Pinterest is a visual platform at its core, so understanding and responding to images is paramount. A few years ago, the content quality team designed and implemented his own batch pipeline that detects similar images. Similarity signals are used extensively on Pinterest for use cases, ranging from improving recommendations based on similar images to removing he spam and abusive content. With over 300 billion Pins, Pinterest offers a unique and growing data set of interests, ideas and intentions. As a data-driven company, Pinterest uses data insights and analytics to make product decisions and reviews to improve his Pinner experience for over 450 million monthly users. I used Pinterest for the recommendation system and image quality for almost every article.
- **Instagram:** Instagram is an American photo and video sharing social networking service owned by Facebook, created by Kevin Systrom and Mike Krieger and first launched on iOS in October 2010. a Fire OS app in June 2014, and an app for Windows 10 in in October 2016. . This can be edited with filters and organized with hashtags and geo-markers. Posts can be made public or shared with his pre-approved followers. User can search other user's content

by tag and location and view trending his content. Users can like photos and follow other users to add their own content to their feed. Instagram does not allow you to download images directly to your system, so you have to take his screenshot of the image and manually save it to the desired folder but Instagram has a lot of high-quality images that can improve the accuracy of your model.

- **Camera:** Last but not least, I took some photos of the food directly from the cell phone's camera. The quality of is not very good, but it can improve the training of the model.



Figure 3.3: Source of data

3.2.2 DATA PRE-PROCESSING

Data pre-processing is an important step in the data mining process. The phrase "garbage in, garbage out" is particularly applicable to data mining, machine learning, and deep learning projects. The way data is collected is often loosely controlled, resulting in out-of-range values, impossible data combinations, and missing values. Analysing data that has not been carefully checked for such issues can lead to misleading results. Therefore, data presentation and quality are of utmost importance before performing analysis. Data pre-processing is often the most important phase of a deep learning project, especially in the computing environment. Knowledge discovery during the training phase becomes more difficult when there is a lot of irrelevant, redundant, noisy, and unreliable data. The data preparation and filtering steps can take significant processing time. Data pre-processing includes cleaning, instance selection, normalization, transformation, feature extraction and selection, etc. The result of pre-processing data is the final training set. Data pre-processing can affect how the final data processing results are interpreted. If the interpretation of the results is a significant issue, this aspect should be carefully considered. As previously mentioned, a total of 15149 images were

collected from all sources. After images were collected, manual filtering was performed to ensure that there were no irrelevant images.

- **Duplicate images:** Duplication in the training set can cause the model to learn a bias against duplicate samples, making the model difficult to generalize to new data. Duplication in the test set leads to inaccurate model performance evaluation and poor performance on hidden data. Manually finding duplicate images in a dataset containing millions of images is an expensive task. You may have downloaded too many images from different sources, resulting in a lot of duplication. In this case, you need a fast, easy-to-use program that finds and removes all duplicates. Therefore, all images were passed to an application called ‘VisiPics’ and checked for duplicates. ‘VisiPics’ doesn’t just look for identical files, it looks for similar images beyond checksum, all in a simple interface. First select the root folder to search and catalogue all images. Then apply five image comparison filters to measure how close the image pairs are. This is a program that is significantly faster than any other commercial product and has an interface that allows you to do other things while automatically finding duplicate images. Recognize two different resolution files of the same image as duplicates, save the same image in different formats, or recognize duplicates with only minor cosmetic changes. All duplicates found are displayed side-by-side with relevant information such as file name, type and size. Auto-select mode can keep high-resolution images, space-saving file types, file sizes smaller than, or all of the above. In this case, the duplicate image threshold is set around the base level of, which is about 60 if the application finds two or more images that are more than or 60 will classify them as duplicates.
- **Data overfitting and underfitting:** Training a deep neural network that can generalize well to new data is a difficult problem. A model with too little capacity cannot learn the problem, while a model with too much capacity can learn too much of the problem and overfit the training set. In both cases, is a model that does not generalize well. A modern approach to reduce generalization error is to use larger models. This may be necessary to use regularization during training that keeps the model weights small. These techniques not only reduce overfitting, but also lead to faster model optimization and better overall performance. I cannot learn the problem regardless of the specific examples in the training data. The model overfits the training data, and new hidden samples and statistical noise added to the samples in the training data set change performance significantly. You might have too many images downloaded from different sources in different classes, so you might have more images in one class and less in another class. Therefore, to create the

best model, we need to define a set of cut-offs for each class to ensure that there is no class imbalance. The range of each Indian dish is kept from a minimum of 250 images to a maximum of 400 images.

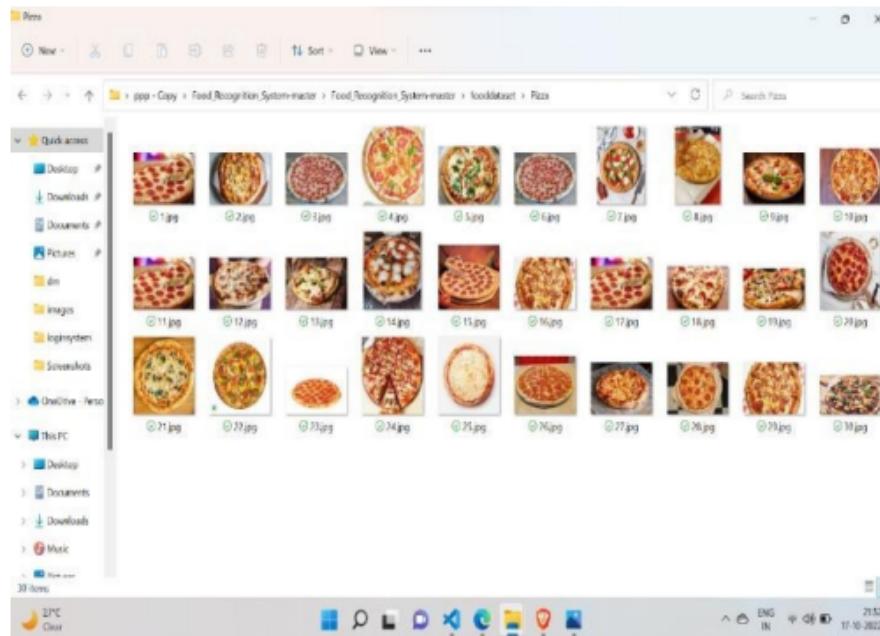


Figure 3.4: DataSet

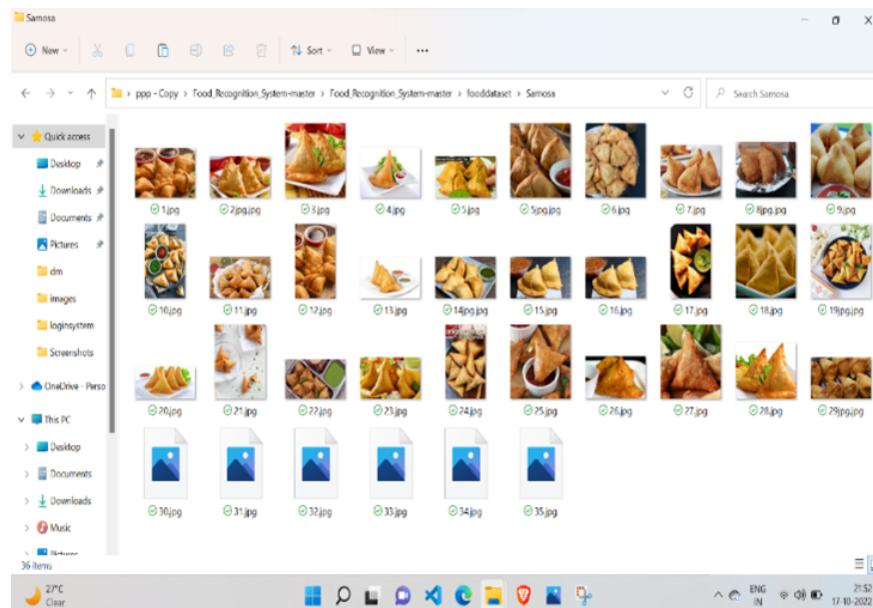


Figure 3.5: DataSet

3.2.3 DATA LABELLING

The next step after data collection and preprocessing is the labeling of selected data. Data labeling is a key step in supervised machine and deep learning tasks. Garbage In Garbage Out is a term commonly used in the deep learning community, meaning that the quality of

the training data determines the quality of the model. The same applies to annotations used to label data. For example, if you show a child a tomato and say it's a potato, the next time your child sees a tomato, he's more likely to recognize it as a potato. Similarly, deep learning models learn by looking at examples, so model results depend on the labels you input during the training phase. Data labeling is a highly manual task. It would be pretty easy if I could find a good open record for the project labeled. So, you have to do the data annotation process yourself. There are several different types of annotations such as bounding box, polygon segmentation, semantic segmentation, 3D boxes, key points and landmarks. In this case, I used the bounding box method for image annotation. The bounding box is a rectangular frame used to define the position of the target object. They can be determined by the axis coordinates x and y of the upper left corner and the axis coordinates x and y of the lower right corner of rectangle. A bounding box annotator can sketch a specific object inside the box as according to the specific requirements of a machine learning or deep learning project.

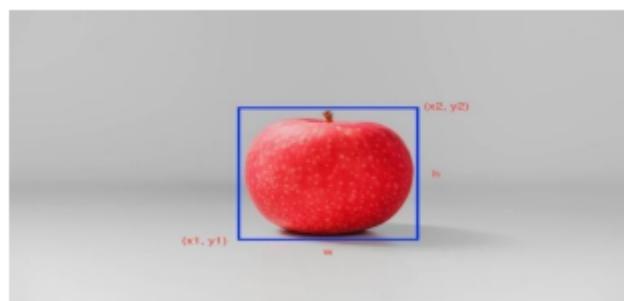


Figure 3.6: Bounding box

3.3 MATHEMATICAL MODEL

- TensorFlow is one of the most widely used deep learning- based object detection methods.
- A detection algorithm not only predict class labels but also detects locations of objects in the input image.
- TensorFlow applies a single neural network to the full image which means that it divides the image into small regions and predicts bounding boxes and probabilities for each region.
- TensorFlow includes a special image recognition function and these images are stored in a specific folder.
- Since the images are relatively equal, implementing this logic for security purposes is trivial.
- It uses Mobilnet model for classification and detection

3.4 MOBILENET

MobileNets are a family of mobile-first computer vision models for TensorFlow, designed to effectively maximize accuracy while addressing the limited resources of on-device or embedded applications. As its name suggests, the MobileNet model is designed for use in mobile applications and is TensorFlow's first mobile computer vision model. MobileNet uses depth separable folds. The number of parameters is greatly reduced compared to networks with regular folds of the same depth in the net. This enables lightweight deep neural networks. Mobilenets are models that perform the same convolutions as CNNs to filter images, but in a different way than previous CNNs. It uses the idea of depth convolution and point convolution which is different from the normal convolution done in normal CNN. This improves the efficiency of CNNs in predicting images and makes them competitive even on mobile systems. We use these types of convolutions as image recognition models because they give a much better answer in a very short time, as they greatly reduce the comparison and recognition time. TensorFlow offers a variety of pretrained models, including: Drag and drop models to identify approximately 1,000 standard objects. Comparison with other similar models like on the Inception model dataset, MobileNet performs better in latency, size, and accuracy. As for the output, the fullfledged model has a significant lag. However, the trade-off is acceptable if the model can be deployed to mobile devices for real-time offline detection.

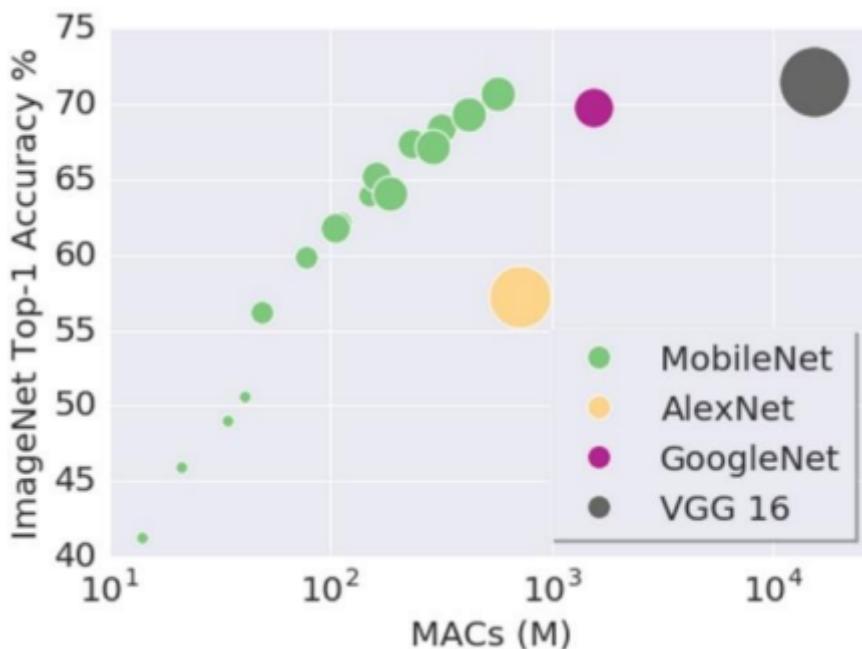


Figure 3.7: Accuracy (y-axis) vs. the number of required operations (x-axis) for available configurations

Chapter 4

EXPERIMENTAL STUDY

4.1 EXPERIMENTAL SETUP

The images used to create the model were downloaded from various sources as described in Section Materials and Methods Data Collection. A total of 4000 images were selected and were transferred for data preprocessing. Finally, all 4000 images acquired after preprocessing were provided as input for model training. We limited the number of images for all 40 classes to avoid overfitting or underfitting the model. The coding part was done in Python language. Python is a programming language that allows you to work faster and integrate systems more effectively. He is strong, quick on his feet, plays well with others, goes everywhere, is friendly, outgoing and very easy to learn. The system model used is an HP Pavilion notebook. RAM is 8192MB. The operating system is Windows 10 Home Single Language 64-bit (10.0, Build 18362).

4.2 PERFORMANCE MEASURES

Object detection models are typically trained on a fixed set of classes, so the model finds and classifies only those classes in images. Also, the object's position is usually in the form of a bounding box. A metric that indicates the accuracy of a particular bounding box is the Intersection over Union (IoU). IoU is the ratio between the intersection and sum of the predicted box and the ground truth box. IoU is used to get true positives and false positives. An IoU ≥ 0.5 is considered a true positive, otherwise it is considered a false positive. Precision and recall are now calculated. IOU is a simple geometric metric that can be easily standardized to assess mAP (mean accuracy) based on a fixed 50 performance, we continually cognizance at the average accuracy of consecutive photographs, mAP, and the common item detection processing speed per second (frames according to 2d, FPS). concerning accuracy, there are numerous techniques used to evaluate the accuracy of item detection fashions or algorithms, but mAP is the primary one. As for speed, FPS is the norm. The performance matrix in this case is used in most object detection algorithms, usually he was

mAP used in TensorFlow. mAP compares the ground truth bounding boxes with the detected boxes and returns a score.

- **Confidence Score:** It reflects the probability that an anchor box contains an object. It is usually predicted by a classifier.
- **Ground Truth Bounding Box:** represents the desired output of an algorithm on an input, for example, the hand labelled bounding box from the testing set that specify where the objects are in the image.
- **Predicted bounding box $x(B_p)$:** represents a rectangle region generated from model detector that indicates the location of the object predicted.
- **Intersection over Union(IoU):** :an evaluation metric used to measure the area encompassed by both the ground-truth bounding box (B_{gt}) the predicted bounding box (B_p).

$$\text{IoU} = \frac{\text{Area of Overlap } (B_{gt} \cap B_p)}{\text{Area of Union } (B_{gt} \cup B_p)}$$

Figure 4.1: IoU Explanation

- **Threshold:** we predefine a threshold of IoU (for instance, 0.5) in classifying whether the prediction is a true positive or a false positive.
- **True Positive (TP):** : A true positive test result is one that detects the condition when the condition is present.
- **True Negative (TN) :** A true negative test result is one that does not detect the condition when the condition is absent.
- **False Positive (FP):** A false positive test result is one that detects the condition when the condition is absent.
- **False Negative(FN) :** A false negative test result is one that does not detect the condition when the condition is present.
- **Precision:** The number of true positive divided by the sum of true positive and false positive.

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP})$$

- **Recall:** the number of true positives divided by the sum of true positives and false negatives. However, the sum is just the number of ground-truths, so it is not necessary to count the number of false negatives.

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN})$$

- **Average precision (AP):** It is not easy to compare the accuracy among different detectors when PR curves intersect with each other. This is why the average precision (AP), based on the precision-recall curve, takes the responsibility for the measurement of object detection accuracy. The average precision's value always falls between 0 and 1 based on the PR curve since Precision and recall are always between 0 and 1, too.
- **Mean average precision (mAP):** The calculation of AP only involves one class. However, in object detection there are usually $K \geq 1$ classes. Mean average precision (mAP) is defined as the mean of AP across all K classes.

$$map = \sum^k AP_i/k$$

Chapter 5

OVERVIEW OF AN APPLICATION

- The user interface will be a web application wherein on opening the application, it will show him a dialog box asking for name of food items he/she is allergic from. Note that this list can be changed afterwards. On the first screen there would be a button to allow user to select the picture he wants to scan. He can select a pre stored image.
- After selecting the image, it will be displayed over the image view. Also, there is a search bar where users can directly look for the food item by selecting its name from the list. Then after clicking on the scan button, the custom computer vision model will be invoked. This model will detect the objects in the image and classify them as per the classes on which it has been trained. The class name of the detected object along with the detected image with a bounding box around the object will be displayed to the user on the screen. In addition to that an estimate total calorie value of all the detected items will be calculated and shown to the user.
- On clicking the Go button the user is then directed to a new page in which the user will find information related to the detected item with name, ingredients, YouTube link for the recipe, nutritional value chart and some facts about the respective food item. On clicking on the name of the detected item, the name is being pronounced using voice integration. Also, if any of the item, user has added to his allergic item's list, is there in the ingredient list, the application will show a red sign stating he is not advised to have that particular food item.
- The application is also integrated with Google Maps and hence if the user wishes to visit any particular restaurant or nearby restaurants, they can just click on the nearby restaurant button and they will be directed to the Google Maps.
- There is also an option for users to look for nearby restaurants serving the specific detected item. There is an options menu with two different options, for first there is an allergies tab

where users can update the items, they are allergic from and for another there is an item tab showing a list of all the available food items stored in the application's database. Hence, this makes my application (Curious Bite) a stand out from other applications present.

- User Interface: These app has a user-friendly interface that allows users to upload an image of their meal and view the results of the object detection model. The interface could also include features such as showing calories to the user and basic value and it's easy to use
- Object Detection Model: The core of the application would be an object detection model trained on a large dataset of food images. The model would be able to recognize different types of food and provide a label for each item detected in the image.

5.1 UNIQUE FEATURES

- The model is designed using object detection techniques so that it will detect different Indian food items and then integrated into an web application.
- The application is set to 40 different food items as for now.
- The application will also provide users with nutritional values of the food items.
- Gallery Integration which will help users to directly scan a previously captured photo from their phone's gallery so that they don't have to click a new photo every time.
- It ensure you the food you are eating is safe for you

5.2 LIMITATIONS

- The application can't detect multiple objects at the same time.
- Sometimes if the images are blur so may be it can't predict proper food item
- The application requires a stable internet connection for smooth functioning.
- Sometimes may be it show inaccurate result

5.3 TECHNOLOGIES USED:

- IDE:- Visual Studio Code
- Languages:-HTML,CSS,PYTHON etc
- Libraries:- Tensorflow,Opencv,NumPy

5.4 DATASET COLLECTION

- Web Scraping
- Google's Open Images
- Pinterest
- Camera
- In summary, cameras are essential devices for capturing images, Pinterest is a popular social media platform for sharing and discovering images and ideas, Google provides access to a vast amount of information and images, and web scraping is a powerful tool for automatically collecting data and images from websites.

5.5 IMPORTANCE OF TENSORFLOW

- TensorFlow is an open-source software library developed by Google for building and training machine learning models. It was released in 2015 and has since become one of the most popular libraries for deep learning. TensorFlow is designed to work with large-scale neural networks and other machine learning algorithms and is widely used for a variety of applications such as computer vision, natural language processing, speech recognition, and robotics.
- TensorFlow is a popular tool for image processing tasks such as image classification, object detection, and segmentation. Here are some ways TensorFlow can be used for image processing:
 - In summary, TensorFlow can be used for a variety of image processing tasks such as image classification, object detection, segmentation, generation, and more. TensorFlow provides a powerful and flexible platform for building and training deep learning models for image processing tasks.



Figure 5.1: Login Page

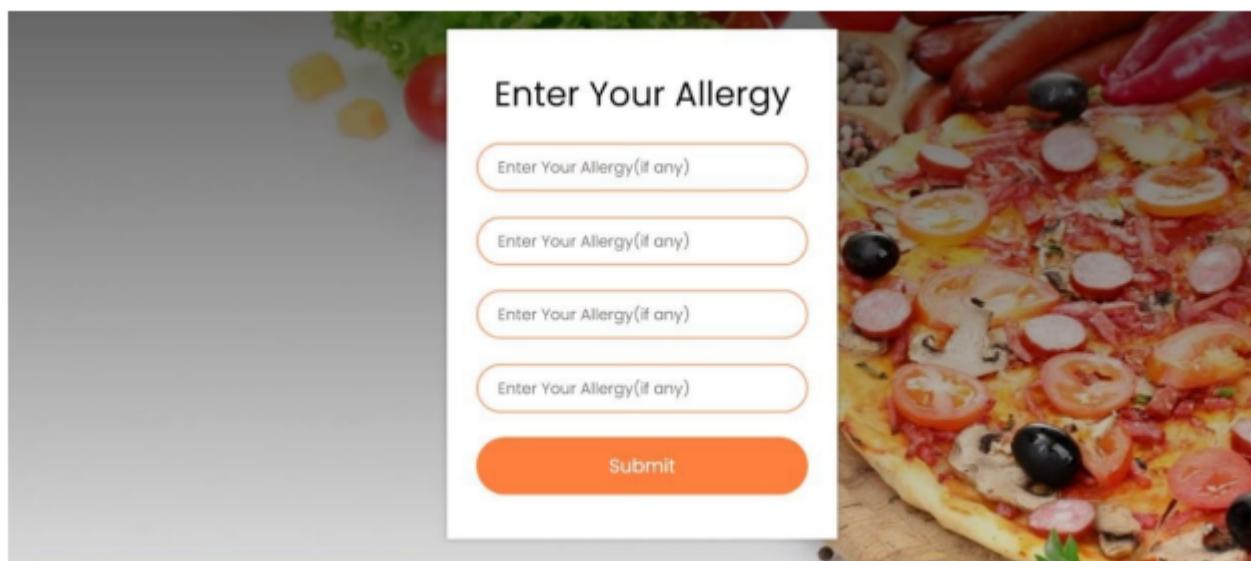


Figure 5.2: Allergy Page

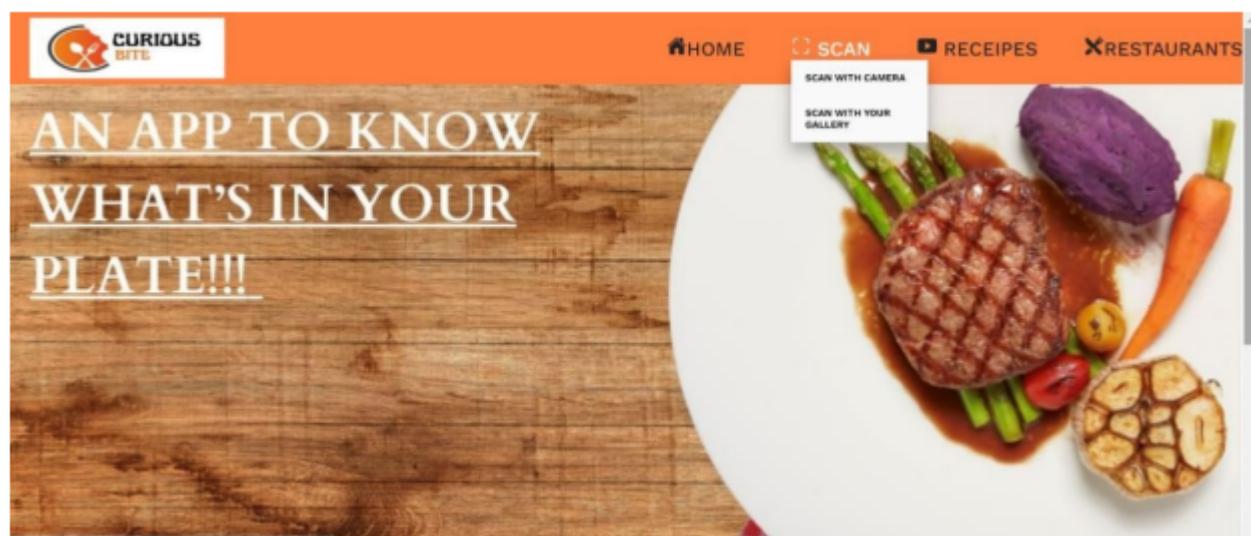


Figure 5.3: Home Page

Food Classifier

Choose Food...



Predict Food

(a)

Food Classifier

Choose Food...



Result: Samosa 🥟

(b)

Figure 5.4: (a) Output (b) Output

Food Classifier

Choose Food...



Predict Food

(a)

Food Classifier

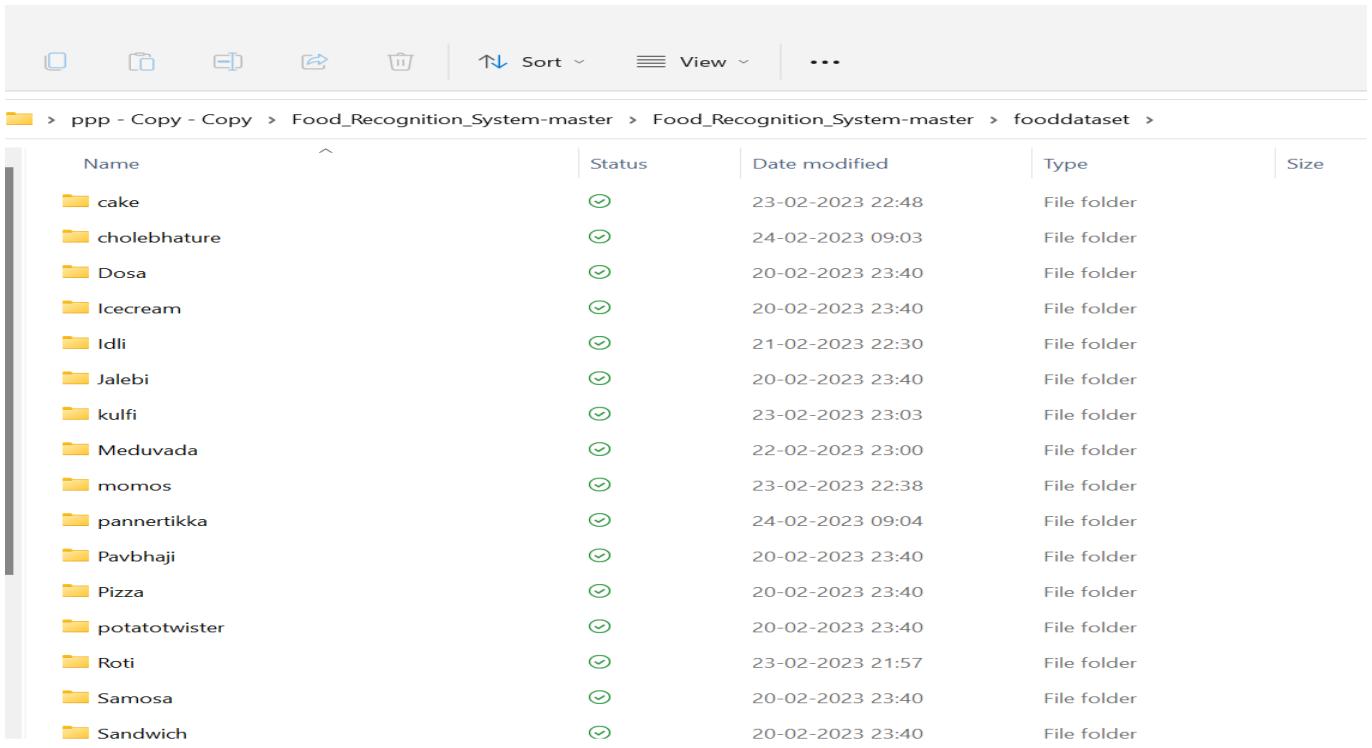
Choose Food...



Result: Pizza 🍕

(b)

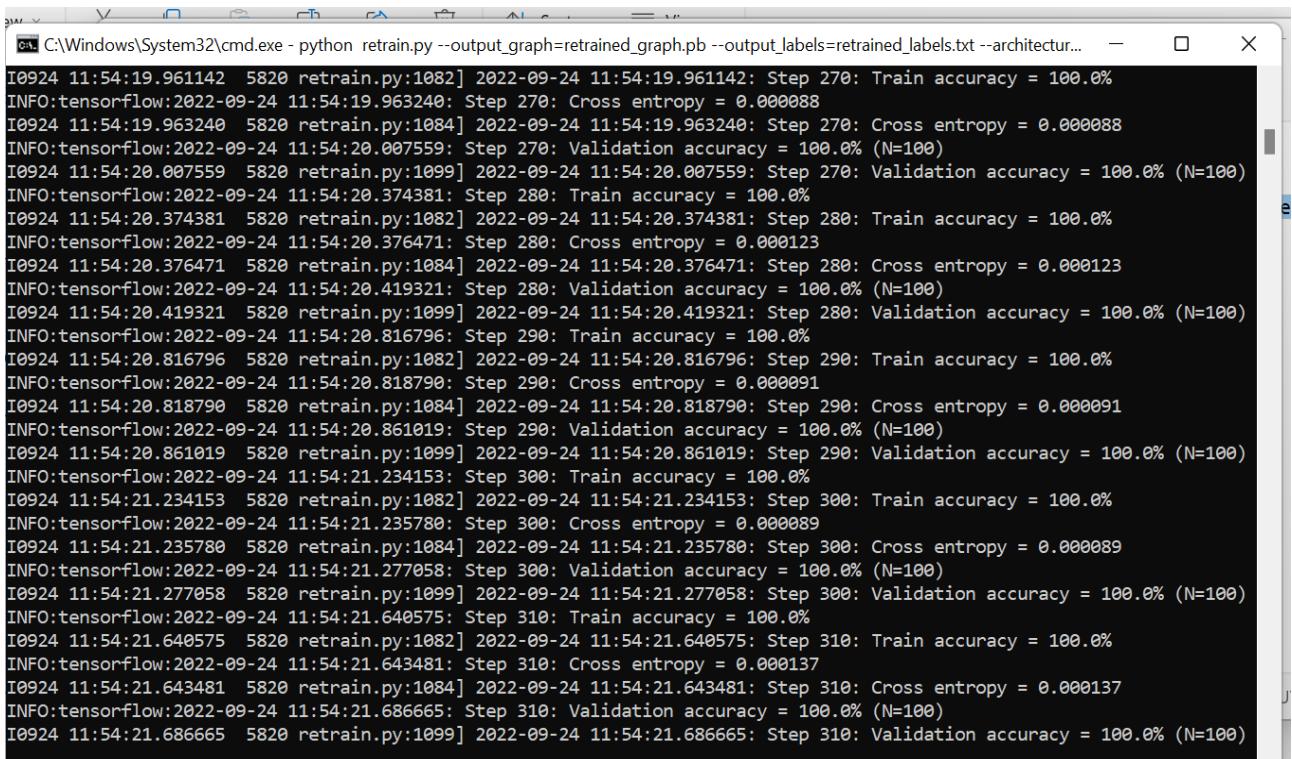
Figure 5.5: (a) Output (b) Output



The screenshot shows a Windows File Explorer window. The path is: ppp - Copy - Copy > Food_Recognition_System-master > Food_Recognition_System-master > fooddataset >. The table lists 15 subfolders under 'fooddataset' with their names, status, date modified, type, and size.

Name	Status	Date modified	Type	Size
cake	✓	23-02-2023 22:48	File folder	
cholebhature	✓	24-02-2023 09:03	File folder	
Dosa	✓	20-02-2023 23:40	File folder	
Icecream	✓	20-02-2023 23:40	File folder	
Idli	✓	21-02-2023 22:30	File folder	
Jalebi	✓	20-02-2023 23:40	File folder	
kulfi	✓	23-02-2023 23:03	File folder	
Meduvada	✓	22-02-2023 23:00	File folder	
momos	✓	23-02-2023 22:38	File folder	
pannertikka	✓	24-02-2023 09:04	File folder	
Pavbhaji	✓	20-02-2023 23:40	File folder	
Pizza	✓	20-02-2023 23:40	File folder	
potatotwister	✓	20-02-2023 23:40	File folder	
Roti	✓	23-02-2023 21:57	File folder	
Samosa	✓	20-02-2023 23:40	File folder	
Sandwich	✓	20-02-2023 23:40	File folder	

Figure 5.6: Food Dataset



```
C:\Windows\System32\cmd.exe - python retrain.py --output_graph=retrained_graph.pb --output_labels=retrained_labels.txt --architectur...
I0924 11:54:19.961142 5820 retrain.py:1082] 2022-09-24 11:54:19.961142: Step 270: Train accuracy = 100.0%
INFO:tensorflow:2022-09-24 11:54:19.963240: Step 270: Cross entropy = 0.000088
I0924 11:54:19.963240 5820 retrain.py:1084] 2022-09-24 11:54:19.963240: Step 270: Cross entropy = 0.000088
INFO:tensorflow:2022-09-24 11:54:20.007559: Step 270: Validation accuracy = 100.0% (N=100)
I0924 11:54:20.007559 5820 retrain.py:1099] 2022-09-24 11:54:20.007559: Step 270: Validation accuracy = 100.0% (N=100)
INFO:tensorflow:2022-09-24 11:54:20.374381: Step 280: Train accuracy = 100.0%
I0924 11:54:20.374381 5820 retrain.py:1082] 2022-09-24 11:54:20.374381: Step 280: Train accuracy = 100.0%
INFO:tensorflow:2022-09-24 11:54:20.376471: Step 280: Cross entropy = 0.000123
I0924 11:54:20.376471 5820 retrain.py:1084] 2022-09-24 11:54:20.376471: Step 280: Cross entropy = 0.000123
INFO:tensorflow:2022-09-24 11:54:20.419321: Step 280: Validation accuracy = 100.0% (N=100)
I0924 11:54:20.419321 5820 retrain.py:1099] 2022-09-24 11:54:20.419321: Step 280: Validation accuracy = 100.0% (N=100)
INFO:tensorflow:2022-09-24 11:54:20.816796: Step 290: Train accuracy = 100.0%
I0924 11:54:20.816796 5820 retrain.py:1082] 2022-09-24 11:54:20.816796: Step 290: Train accuracy = 100.0%
INFO:tensorflow:2022-09-24 11:54:20.818790: Step 290: Cross entropy = 0.000091
I0924 11:54:20.818790 5820 retrain.py:1084] 2022-09-24 11:54:20.818790: Step 290: Cross entropy = 0.000091
INFO:tensorflow:2022-09-24 11:54:20.861019: Step 290: Validation accuracy = 100.0% (N=100)
I0924 11:54:20.861019 5820 retrain.py:1099] 2022-09-24 11:54:20.861019: Step 290: Validation accuracy = 100.0% (N=100)
INFO:tensorflow:2022-09-24 11:54:21.234153: Step 300: Train accuracy = 100.0%
I0924 11:54:21.234153 5820 retrain.py:1082] 2022-09-24 11:54:21.234153: Step 300: Train accuracy = 100.0%
INFO:tensorflow:2022-09-24 11:54:21.235780: Step 300: Cross entropy = 0.000089
I0924 11:54:21.235780 5820 retrain.py:1084] 2022-09-24 11:54:21.235780: Step 300: Cross entropy = 0.000089
INFO:tensorflow:2022-09-24 11:54:21.277058: Step 300: Validation accuracy = 100.0% (N=100)
I0924 11:54:21.277058 5820 retrain.py:1099] 2022-09-24 11:54:21.277058: Step 300: Validation accuracy = 100.0% (N=100)
INFO:tensorflow:2022-09-24 11:54:21.640575: Step 310: Train accuracy = 100.0%
I0924 11:54:21.640575 5820 retrain.py:1082] 2022-09-24 11:54:21.640575: Step 310: Train accuracy = 100.0%
INFO:tensorflow:2022-09-24 11:54:21.643481: Step 310: Cross entropy = 0.000137
I0924 11:54:21.643481 5820 retrain.py:1084] 2022-09-24 11:54:21.643481: Step 310: Cross entropy = 0.000137
INFO:tensorflow:2022-09-24 11:54:21.686665: Step 310: Validation accuracy = 100.0% (N=100)
I0924 11:54:21.686665 5820 retrain.py:1099] 2022-09-24 11:54:21.686665: Step 310: Validation accuracy = 100.0% (N=100)
```

Figure 5.7: Image Training

Chapter 6

CONCLUSION & FUTURE WORK

- In this project, we have trained a CNN object detection model (TensorFlow) at desktop platform.
- As a baseline, I have a running Web application that runs my custom object detection model trained by TensorFlow.
- Many unique features are also present in my application such as route optimization, Google integration and YouTube integration.
- There are only 40 different food items added in the application as for now. Maybe, in the next update we will add a few more items for the users to help them get more and more information as possible.
- In addition to that I am planning to add more images with different sharpness, brightness, rotation and blurred images for better performance of the model. Also, more features like diet tips based on user's body ratio and tracking what the user is eating throughout the day can be added at later stages.
- Things which we are going to implement: -
 - We will show some additional nutritional facts about the detected or searched item.
 - Also, we will add feature of voice recognition i.e., it can help users to easily pronounce the name of detected item.
 - Users can add list of food items that they are allergic from and if the item they searched have the respective ingredient, the app will show a red sign stating the user that he is not advised to eat that item.

REFERENCES

- [1] Talele, Ajay, Aseem Patil, and Bhushan Barse. "Detection of real time objects using TensorFlow and OpenCV." Asian Journal For Convergence In Technology (AJCT) ISSN-2350-1146 (2019).
- [2] Li, Meian Janahiraman, Tiagrajah V., and Mohamed Shahrul Mohamed Subuhan."Traffic light detection using tensorflow object detection framework." 2019 IEEE 9th International Conference on System Engineering and Technology (ICSET). IEEE, 2019.
- [3] Gavai, Nitin R., et al. "MobileNets for flower classification using TensorFlow." 2017 international conference on big data, IoT and data science (BID). IEEE, 2017.
- [4] Raj, Rishu, Soumya Michael, and D. Femi. "Forest Monitoring System through RFID and Tensorflow Object Detection." 2019 International Carnahan Conference on Security Technology (ICCST). IEEE, 2019.
- [5] Fu, Kangwei, et al. "Text detection for natural scene based on mobilenet V2 and U-net." 2019 IEEE international conference on mechatronics and automation (ICMA). IEEE, 2019.
- [6] Ege, T., Yanai, K. (2017, November). Estimating food calories for multiple-dish food photos. In 2017 4th IAPR Asian Conference on Pattern Recognition (ACPR) (pp. 646-651). IEEE.
- [7] Zhang, W., Zhao, D., Gong, W., Li, Z., Lu, Q., Yang, S. (2015, August). Food image recognition with convolutional neural networks. In 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC- ATC-ScalCom) (pp. 690-693). IEEE.
- [8] Taskiran, M., Kahraman, N. (2019, July). Comparison of CNN Tolerances to Intra Class Variety in Food Recognition. In 2019 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) (pp. 1-5). IEEE.

- [9] Fonseca, L. M. G., Namikawa, L. M., Castejon, E. F. (2009, October). Digital image processing in remote sensing. In 2009 Tutorials of the XXII Brazilian Symposium on Computer Graphics and Image Processing (pp. 59-71). IEEE.
- [10] Zhang, H., Wang, L., Jia, R., Li, J. (2009, October). A distance measuring method using visual image processing. In 2009 2nd International Congress on Image and Signal Processing (pp. 1-5). IEEE.
- [11] Akbar, J., Shahzad, M., Malik, M. I., Ul-Hasan, A., Shafait, F. (2019, December). Runway detection and localization in aerial images using deep learning. In 2019 Digital Image Computing: Techniques and Applications (DICTA) (pp. 1-8). IEEE.
- [12] Tiankaew, U., Chunpongthong, P., Mettanant, V. (2018, July). A Food Photography App with Image Recognition for Thai Food. In 2018 Seventh ICT International Student Project Conference (ICT-ISPC) (pp. 1-6). IEEE.
- [13] Chin, C. L., Huang, C. C., Lin, B. J., Wu, G. R., Weng, T. C., Chen, H. F. (2016, November). Smartphone-based food category and nutrition quantity recognition in food image with deep learning algorithm. In 2016 International Conference on Fuzzy Theory and Its Applications (iFuzzy) (pp. 1-1). IEEE.
- [14] Lasod, A., Soni, D. (2017, August). Efficiency enhancement of food recognition using artificial neural network. In 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS) (pp. 2758-2762). IEEE.
- [15] hin, C. L., Huang, C. C., Lin, B. J., Wu, G. R., Weng, T. C., Chen, H. F. (2016, November). Smartphone-based food category and nutrition quantity recognition in food image with deep learning algorithm. In 2016 International Conference on Fuzzy Theory and Its Applications (iFuzzy) (pp. 1-1). IEEE.
- [16] Arcila, O., Dinas, S., Bañón, J. M. (2012, October). Collision detection model based on Bounding and containing Boxes. In 2012 XXXVIII Conferencia Latinoamericana En Informatica (CLEI) (pp. 1-10). IEEE.
- [17] Tu, Y. L., Lin, W. Y., Lin, Y. C. (2020, November). Automatic Leaf Counting Using Improved yolov3. In 2020 International Symposium on Computer, Consumer and Control (IS3C) (pp. 197- 200). IEEE.
- [18] E.Makinenand R.Raisamo. Evaluation of gender classification methods with automatically detected and aligned faces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(3):541–547, March 2008

- [19] S. Yeom, D. Lee, Y. Jang, M. Lee, and S. Jung, "Real-time concealed-object detection and recognition with passive millimeter wave imaging," *Optics Express*, Vol. 20, No. 9, pp. 9371-9381 (2012).
- [20] Chauhan, N. S., Rahman, F., Sarker, R., Pious, M. M. H.(2018). Vehicle detection, tracking and counting using linear quadratic estimation technique. 2018 2nd International Conference on Inventive Systems and Control (ICISC).
- [21] van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R., Schatz, M. (1996). A new generation of the IMAGIC image processing system. *Journal of structural biology*, 116(1), 17-24.
- [22] Du, C. J., Sun, D. W. (2004). Recent developments in the applications of image processing techniques for food quality evaluation. *Trends in food science technology*, 15(5), 230-249.
- [23] Jackman, P., Sun, D. W. (2013). Recent advances in image processing using image texture features for food quality assessment. *Trends in food science technology*, 29(1), 35-43.
- [24] Hassannejad, H., Matrella, G., Ciampolini, P., De Munari, I., Mordonini, M., Cagnoni, S. (2016, October). Food image recognition using very deep convolutional networks. In *Proceedings of the 2nd international workshop on multimedia assisted dietary management* (pp. 41-49).
- [25] Joshua, S. R., Shin, S., Lee, J. H., Kim, S. K. (2023). Health to Eat: A Smart Plate with Food Recognition, Classification, and Weight Measurement for Type-2 Diabetic Mellitus Patients' Nutrition Control. *Sensors*, 23(3), 1656.
- [26] Amit, Y., Felzenszwalb, P., Girshick, R. (2020). Object detection. *Computer Vision: A Reference Guide*, 1-9.
- [27] Cai, Q., Li, J., Li, H., Weng, Y. (2019, February). BTBUFood-60: Dataset for object detection in food field. In *2019 IEEE International Conference on Big Data and Smart Computing (BigComp)* (pp. 1-4). IEEE.
- [28] Zhou, L., Zhang, C., Liu, F., Qiu, Z., He, Y. (2019). Application of deep learning in food: a review. *Comprehensive reviews in food science and food safety*, 18(6), 1793-1811.
- [29] Bishop, C. M. (2021). Neural networks and their applications. *Review of scientific instruments*, 65(6), 1803-1832.
- [30] Gupta, N. (2013). Artificial neural network. *Network and Complex Systems*, 3(1), 24-28.
- [31] Zhang, Z., Zhang, Z. (2018). Artificial neural network. *Multivariate time series analysis in climate and environmental research*, 1-35.

- [32] Sainz-De-Abajo, B., García-Alonso, J. M., Berrocal-Olmeda, J. J., Laso-Mangas, S., De La Torre-Díez, I. (2020). FoodScan: Food monitoring app by scanning the groceries receipts. *IEEE Access*, 8, 227915-227924.
- [33] Knez, S., Šajn, L. (2020). Food object recognition using a mobile device: evaluation of currently implemented systems. *Trends in Food Science Technology*, 99, 460-471.
- [34] Kawano, Y., Yanai, K. (2013). Real-time mobile food recognition system. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 1-7).
- [35] Tanno, R., Okamoto, K., Yanai, K. (2016, October). Deepfoodcam: A dcnn-based real-time mobile food recognition system. In *Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management* (pp. 89-89).
- [36] Albawi, S., Mohammed, T. A., Al-Zawi, S. (2017, August). Understanding of a convolutional neural network. In *2017 international conference on engineering and technology (ICET)* (pp. 1-6). Ieee.
- [37] Pang, B., Nijkamp, E., Wu, Y. N. (2020). Deep learning with tensorflow: A review. *Journal of Educational and Behavioral Statistics*, 45(2), 227-248.
- [38] Ramesh, A., Sivakumar, A., Angel, S. S. (2020, December). Real-time food-object detection and localization for indian cuisines using deep neural networks. In *2020 IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT)* (pp. 1-6). IEEE.
- [39] Rakhsith, L. A., Anusha, K. S., Karthik, B. E., Nithish, D. A., Kumar, V. K. (2021). A survey on object detection methods in deep learning. In *Proc. of 2021 Second Int. Conf. on Electronics and Sustainable Communication Systems (ICESC)*.
- [40] Kasyap, V. B., Jayapandian, N. (2021, May). Food Calorie Estimation using Convolutional Neural Network. In *2021 3rd International Conference on Signal Processing and Communication (ICPSC)* (pp. 666-670). IEEE.

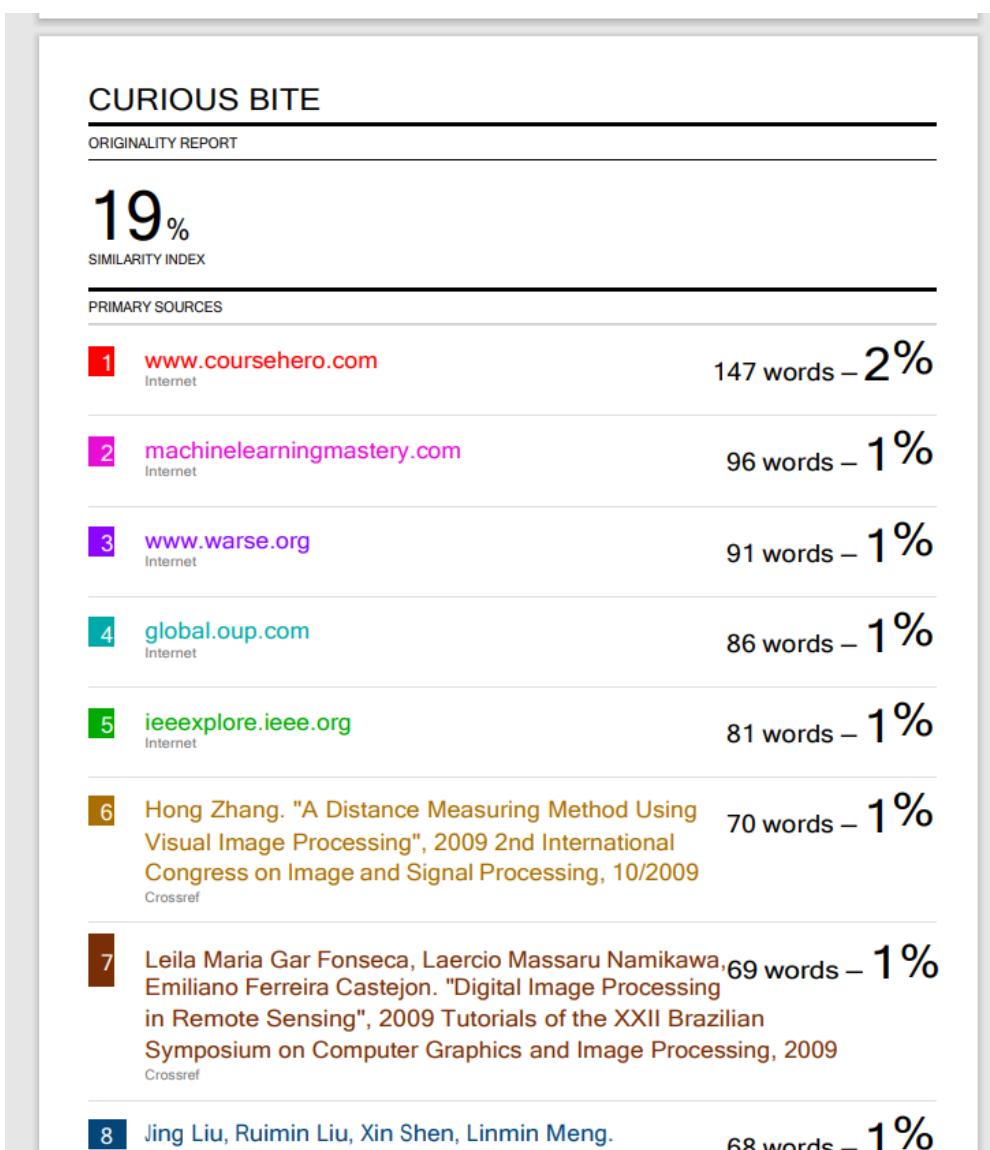


Figure 6.1: ORIGINALITY REPORT