

Bangladeshi Food Detection Using Yolov5

One semester long project report to be submitted in partial fulfilment of the requirements for the degree

of

Bachelor of Science in Computer Science and Engineering

by

Ashikuzzaman Pias
172014012

Under the supervision of

Dr. Mohammed Ashikur Rahman
Assistant Professor



COMPUTER SCIENCE AND ENGINEERING
UNIVERSITY OF LIBERAL ARTS BANGLADESH

SPRING 2022

© *Ashikuzzaman Pias*

All rights reserved

DECLARATION

Project Title Bangladeshi Food Detection Using YOLOv5

Author *Ashikuzzaman Pias*

Student ID 172014012

Supervisor Dr. Mohammed Ashikur Rahman

I declare that this one semester long project report entitled *Bangladeshi Food Detection Using YOLOv5* is the result of my own work except as cited in the references. The project report has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Ashikuzzaman Pias
172014012

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Date: 14th June 2022

CERTIFICATE OF APPROVAL

This is to certify that the one semester long project report entitled **Bangladeshi Food Detection Using YOLOv5**, submitted by **Ashikuzzaman Pias** (Student ID: 172014012) is an undergraduate student of the **Department of Computer Science and Engineering** has been examined. Upon recommendation by the examination committee, we hereby accord our approval of it as the presented work and submitted report fulfil the requirements for its acceptance in partial fulfilment for the degree of *Bachelor of Science in Computer Science and Engineering*.

Dr. Mohammed Ashikur Rahman
Assistant Professor

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Dr. Abul Kalam Al Azad
Associate Professor and Head (Acting)

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Place: Dhaka

Date: 14th June 2022

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my project supervisor, Dr. Mohammed Ashikur Rahman, for giving me the opportunity to conduct research and providing invaluable guidance throughout this work. His dynamism, vision, sincerity and motivation have deeply inspired me. He has taught me the methodology to carry out the work and to present the works as clearly as possible. It was a great privilege and honour to work and study under his guidance.

I am greatly indebted to my honourable teachers of the Department of Computer Science and Engineering at the University of Liberal Arts Bangladesh who taught me during the course of my study. Without any doubt, their teaching and guidance have completely transformed me to the person that I am today.

I am extremely thankful to my parents for their unconditional love, endless prayers and caring, and immense sacrifices for educating and preparing me for my future. I would like to say thanks to my friends and relatives for their kind support and care.

I would like to express gratitude towards my brother who has always been supportive to me and stayed by my side at every situation and also for caring me like his own child.

Finally, I would like to thank all the people who have supported me to complete the project work directly or indirectly.

Ashikuzzaman Pias

University of Liberal Arts Bangladesh

Date: 14th June 2022

To my mother
Piara Begum
and father
Late Abdul Kayum

- **Ashikuzzaman Pias**

ABSTRACT

Bangladesh is a country of having diverse cuisine. Many people from different parts of the world visit Bangladesh. **Bangladeshi cuisine** has is very popular to people of many countries. But many people are unfamiliar with the names of different Bangladeshi dishes. So this is a **Machine Learning Approach** to make the Bangladeshi dishes visible around the world. I have created a **dataset** classifying **21 Bangladeshi dishes** using machine learning. After collecting images for processing the dataset was created. After creating the dataset, the **Yolov5** model was trained on the dataset. The model showed good performance having mAP_0.5 score of 97.5% and mAp_0.5:0.95 score of 92.1%. The model was able to detect all the given classes correctly. This dataset can be useful in various purposes in future. It can be used in calorie estimation, dietary assessment and other purposes. There are future scopes for this project. The model can be used to build an Android Application to detect foods from images and videos. The application can be useful for children to get familiar with different dishes and also for tourists who don't recognise Bangladeshi foods properly.

Keywords: Bangladeshi cuisine, Machine Learning Approach, dataset, 21 Bangladeshi dishes, Yolov5.

Contents

1 Introduction

1.1 Introduction	1
1.2 Motivation	1
1.3 Rationale of the Study	1
1.4 Research Questions	2
1.5 Expected Output	2
1.6 Project Management and Finance	2
1.6.1 Project Timeline	2
1.6.2 Finances	3

2 Background

2.1 Terminologies	4
2.2 Related Works	4
2.3 Comparative Analysis and Summary	5
2.4 Scope of the Problem	5
2.5 Challenges	5

3 Research Methodology

3.1 Research Subject and Instrumentation	7
3.1.1 Research Subject	7
3.1.2 Instrumentation	7
3.2 Data Collection Procedure	7
3.3 Statistical Analysis	8
3.4 Proposed Methodology	10
3.4.1 Architecture of Yolov5	10
3.5 Implementation Requirements	12

4 Experimental Results and Discussion

4.1 Experimental Setup	13
4.2 Experimental Results and Analysis	13
4.2.1 Yolov5s	13
4.2.2 Yolov5m	14

4.2.3 YOLOv5l	15
4.2.4 Comparative Analysis of the training models	15
4.2.5 Precision-Recall curve analysis	16
4.3 Discussion	18
5 Impact on Society, Environment and Sustainability	
5.1 Impact on Society	22
5.2 Impact on Environment	22
5.3 Ethical Aspects	23
6 Summary, Conclusion, Recommendation and Implication for Future Research	
6.1 Summary of the Study	24
6.2 Conclusions	24
6.3 Implication for Further Study	24
References	25

List of Tables

1.6.1 Project Timeline	2
3.1.2 Instrumentation	7
3.3 Statistical Analysis of the Dataset	8
4.2.4 Comparative Analysis of the training models	15

List of Figures

1.6.1	Project Timeline Gantt Chart	3
3.2	Data Collection Procedure	8
3.3	Statistical Analysis of the Dataset	9
3.4	Machine Learning Approach	10
3.4.1	Yolov5 Architecture	11
4.2.1	Yolov5s Evaluation Metrics	13
4.2.2	Yolov5m Evaluation Metrics	14
4.2.3	Yolov5l Evaluation Metrics	15
4.2.5 (1)	Precision-Recall curve for Yolov5s	16
4.2.5 (2)	Precision-Recall curve for Yolov5m	17
4.2.5 (3)	Precision-Recall curve for Yolov5l	17

Chapter 1

Introduction

1.1 Introduction

Bangladesh is a land of tradition and culture. Bangladeshi people have some great habits being followed for hundreds of years. One of them is the Bangladeshi food. Bangladeshi food is very distinctive in world culture. People of this country like to eat and love to make others eat. Bangladeshi cuisine is consisted of a wide range of delicious and popular dishes. Where works related to foods in different countries have a great stage in computer science, but there is not much work done for Bangladeshi foods so far. So in this project it will be an attempt to create a dataset for 21 Bangladeshi dishes and will train that dataset to detect Bangladeshi foods. What is a dataset? Dataset is a collection of data which can be used for different research purposes. Dataset is needed in machine learning systems to perform different studies and actions. Datasets help us to do further researches in different fields. The food dataset is very important for recognising Bangladeshi foods. In this dataset I will process around 5000 images and train the dataset for recognising Bangladeshi foods.

1.2 Motivation

There have been works relating to food detection around the world. A paper was published in 2017 where they achieved 51 % accuracy which is relatively low (Giovany, Putra, Hariawan & Wulandhari, 2017). Suma et al. published a paper in 2020 where they used fewer categories of foods (Suma, Bharathi, Harikumar & Deepak, 2020). There are more projects relating to this field having more issues and also few of them were not inducted for human use. There have been few works done for Bangladeshi food classification. Uddin et al. published a paper where they achieved over 98% accuracy but they classified only seven Bangladeshi foods (Uddin, Miraj, Sarma, Das & Gany, 2021). Jeny et al. did a paper in 2019 where they classified only six Bangladeshi foods (Jeny, Junayed, Ahmed, Habib & Rahman, 2019). Moreover, Bangladesh has some very beautiful tourist destinations. People from different countries around the world visit Bangladesh every year. One of their main attractions in the country is the Bangladeshi food. If the dataset for Bangladeshi food is available, then there might be different research scopes and also people will get to know about Bangladeshi foods more. So this dataset will be helpful in creating a scope for Bangladeshi foods around the world.

1.3 Rationale of the study

The project is named ‘Bangladeshi food detection using Yolov5’. In this project there is a dataset which consists of 5040 images and classifies 21 food categories. Other works in Bangladeshi foods were not much significant during my research, So, this project is designed to provide a better performance than others. A proper research in this area has been carried out before working on this project. All the food categories have been properly labelled to avoid any kind of problem during the development. The system shows good performance during

testing. The proper labelling of the images has a significant contribution in the better performance of the dataset.

1.4 Research Questions

The project is centred to the following question:

1. How to categorize the Bangladeshi foods from image using Machine Learning Technique?

1.5 Expected Output

The outcome of the project will be a proper dataset and a trained model will good accuracy. I am expecting a dataset consisting 5040 images and a Yolov5 custom data trained model with more than 85% accuracy.

1.6 Project Management and Finance

1.6.1 Project Timeline

This project has several portions which need to be completed. The project is divided into different sections and each portion will be completed in a specific time frame. The timeline of this research project is as follows:

Table 1.6.1: Project Timeline

Tasks	Start Date	End Date	Duration (Days)
1. Planning			
1.1. Project selection	4-Mar	6-Mar	2
1.2. Collect images for annotation	7-Mar	4-Apr	28
2. Research			
2.1. Labellimg	5-Apr	7-Apr	2
2.2. Yolov5	8-Apr	12-Apr	4
2.3. Dataset	13-Apr	18-Apr	5
3. Implementation			
3.1. Image Annotation	19-Apr	4-May	15
3.2. Model Train	5-May	15-May	10
4. Report			
4.1. Final Report Writing	16-May	6-Jun	21
4.2. Final Report Submission	7-Jun	9-Jun	2

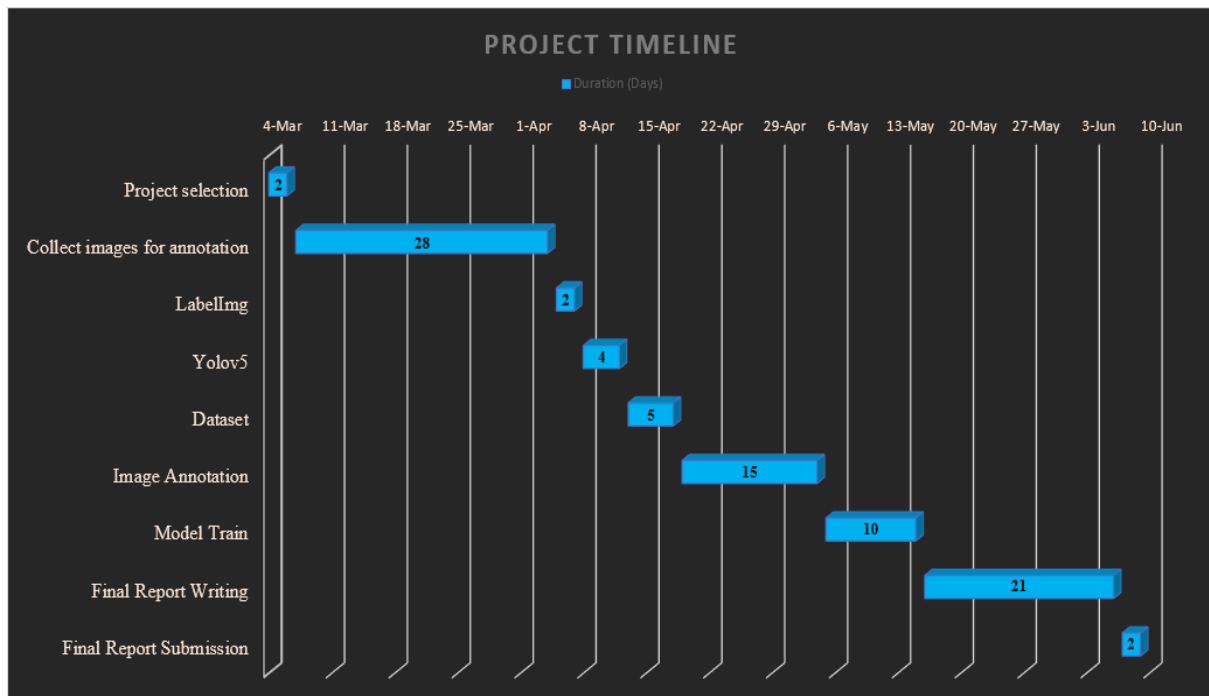


Figure 1.6.1: Project Timeline Gantt Chart

1.6.2 Finances

This research project will be developed solely. The resources to be used are free to use, thus it will result in no financial cost during the research.

Chapter 2

Background

2.1 Terminologies

Yolov5: Yolov5 is a state of the art object detection algorithm. It was developed by Ultralytics in 2020. It is currently the most advanced object detection algorithm available (Garg, 2021).

Dataset: A dataset is a collection of data. In machine learning a dataset consists of data with their corresponding labels. Dataset is regarded as the most important element in machine learning.

Image Processing: Image processing a method of performing some operations on images to enhance the image or get essential information from an image.

2.2 Related Works

There have been works related to food detection around the world. There are some very extensive works which are very successfully completed. There are also different issues with different projects. During my study about this project, I found out some relevant articles connected to my project. A research paper titled “**Traditional Bengali Food Classification Using Convolutional Neural Network**” classified seven Bangladeshi dishes during their research. Their contribution is the development of a dataset of Traditional Bengali Food Image (TBFI). They included images of seven different traditional Bengali foods: **Biriyani, Panta Ilish, Khichuri, Fuchka, Roshogolla, Dim Vuna & Kala Vuna**. For doing their implementation, they developed a scratch model incorporating **Convolutional Neural Network (CNN)**, which rectifies another vital contribution. Because traditional Neural Networks do not perform well on picture datasets, the CNN technique was chosen due to its high accuracy, computing power with efficiency, and automated detection of crucial elements without human intervention. Furthermore, for TBFI classification, a transfer learning strategy with fine-tuned VGG16 was applied. The suggested model in this research produced a final result on their TBFI dataset with an average accuracy of **98%** in categorizing traditional Bengali cuisine photos (Uddin, Miraj, Sarma, Das & Gany, 2021).

In 2019, a paper titled “**FoNet - Local Food Recognition Using Deep Residual Neural Networks**” was published where the researchers had classified six classes of food images. They had created and utilized a Deep Residual Neural Network for grouping the categories. They also employed two convolutional neural networks to differentiate six groups of food photos. The proposed model was then compared to these two deep learning models. Their proposed model has an impressive maximum accuracy of **98.16%** (Jeny, Junayed, Ahmed, Habib & Rahman, 2019).

In 2017, a group of researchers published a study on Indonesian foods where they used SIFT and other techniques named **K-Dimensional Tree (K-D Tree)** and **Backpropagation Neural Network (BPNN)**. They applied these techniques on three types of Indonesian foods namely

Bakso, Ayam bakar and **Sate**. During this approach, they achieved **51%** and **44%** accuracy for **BPNN** and **K-D Tree** respectively (Giovany, Putra, Hariawan & Wulandhari, 2017).

In a research paper titled “**Deep Learning Techniques for Visual Food Recognition on a Mobile App**” the authors used two datasets named UPMC and ETHZ and run a **Convolutional Neural Network (CNN)**. They developed an Android Application after employing the training model. In their application, they had achieved **73%** accuracy (Bonis, Amato, Falchi, Gennaro & Manghi, n.d.).

Convolutional Neural Network (CNN) was used in a study named “**Food Recognition Using Deep Convolutional Neural Network**” to detect and recognize food images as well as assess nutrients in the food. Given the variety of foods, picture identification of food products is particularly difficult. Food-101 dataset was utilized for training and testing the model; 101 distinct classes of food photos were employed to increase model accuracy. The proposed model provided more than **80%** accuracy (Priyanka & Geetha, 2022).

2.3 Comparative Analysis and Summary

While researching in the papers discussed above, I have noticed different approaches by different authors and they had proper justification of the methods they have used. I have noticed the use of **Convolutional Neural Network (CNN)** in most of the papers. The CNN model has mixed number of performance results. I have seen CNN providing **98%** accuracy as well as **73%** accuracy. The datasets are a very important factor in training models. The **BPNN** and **K-D Tree** had relatively low accuracy while detecting objects. All the papers have different contributions in the field of object detection. In this technical era, object detection is one of the most popular research field. So, different authors try to contribute as much as they can in the field of Machine Learning.

2.4 Scope of the Problem

During my research I will contribute a dataset consisting of around 5000 images. I am going to classify 21 Bangladeshi food categories and after processing the images, I will train a Yolov5 model and evaluate the performance of the model on my dataset.

2.5 Challenges

As I am doing this project solely, so there are few challenges that needs to be addressed. The major challenge is related to model training. The challenges during this research include:

1. Data acquisition.
2. Dataset processing.
3. Model Training.

Data acquisition: As this is a single project, so collecting large amount of images is a complex task. Large amount of data sometimes includes junk files. So I need to address those issues and collect images properly.

Dataset processing: Processing 5000 thousand images is a very tedious work. During annotating the images, I need to address that all the images are annotated perfectly and annotated correctly. The bounding boxes cannot be too large or too small. I need to ensure that all the images have their corresponding xml files.

Model Training: I will be using the YOLOv5 training model for my dataset. There are four variants of the YOLOv5 model. The YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. These models have different performance metrics. So I need to ensure which training model I am using for my dataset.

Chapter 3

Research Methodology

3.1 Research Subject and Instrumentation

3.1.1 Research Subject

The project is titled as “**Bangladeshi Food Detection Using YOLOv5**”. In this project, the research subject is object detection. Based on the dataset that I will be developing, the training model will be employed on the dataset and the performance of the model will be evaluated. For my project I have chosen 21 Bangladeshi food categories.

3.1.2 Instrumentation

During the progression of the project, several tools will be used as resources. The tools to be used in this project are given below:

Table 3.1.2: Instrumentation

Name	Purpose
1. LabelImg	Image Processing
2. YOLOv5	Training Model
3. Google Colaboratory	IDE

3.2 Data Collection Procedure

Data collection is the most important part while working with datasets. For my project I have created a dataset consisting of 5040 images. I have collected in two different ways. I have taken pictures and videos of different foods through my mobile phone and also used videos from the internet for collecting the data. I have extracted frames from the videos and stored them as image files. After data collection, it was seen that around 55% of the images are from internet and 45% of the images have been collected physically.

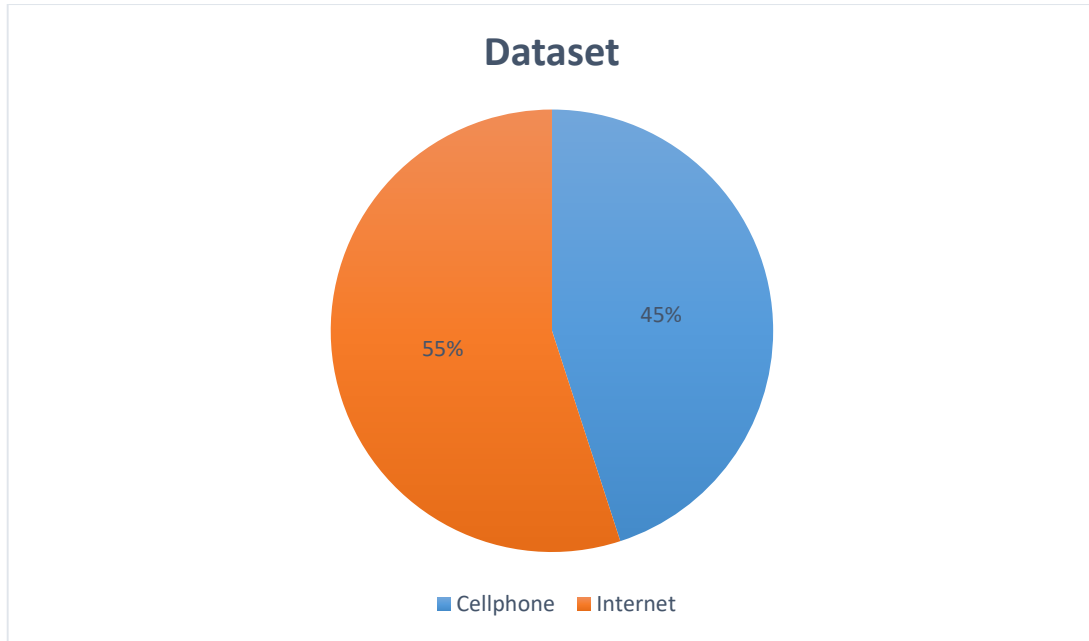


Figure 3.2: Data Collection Procedure

3.3 Statistical Analysis

I have collected 5040 images which include 21 food categories. In the table below, I will represent a statistical figure of all the food categories and their corresponding number of images.

Table 3.3: Statistical Analysis of the Dataset

Category	Number of Images
1. Bhaat (Plain rice)	240
2. Dal (Pulse)	240
3. Aloo vorta (Mashed Potato)	240
4. Khichuri	240
5. Chicken roast	240
6. Chicken curry	240
7. Beef curry	240
8. Biryani	240
9. Jhal muri	240
10. Fuchka	240
11. Puri	240
12. Singara	240
13. Samocha	240

14. Vapa pitha	240
15. Roshogolla	240
16. Chomchom	240
17. Kalo jam	240
18. Jilapi	240
19. Misti doi	240
20. Borhani	240
21. Falooda	240
Total	5040

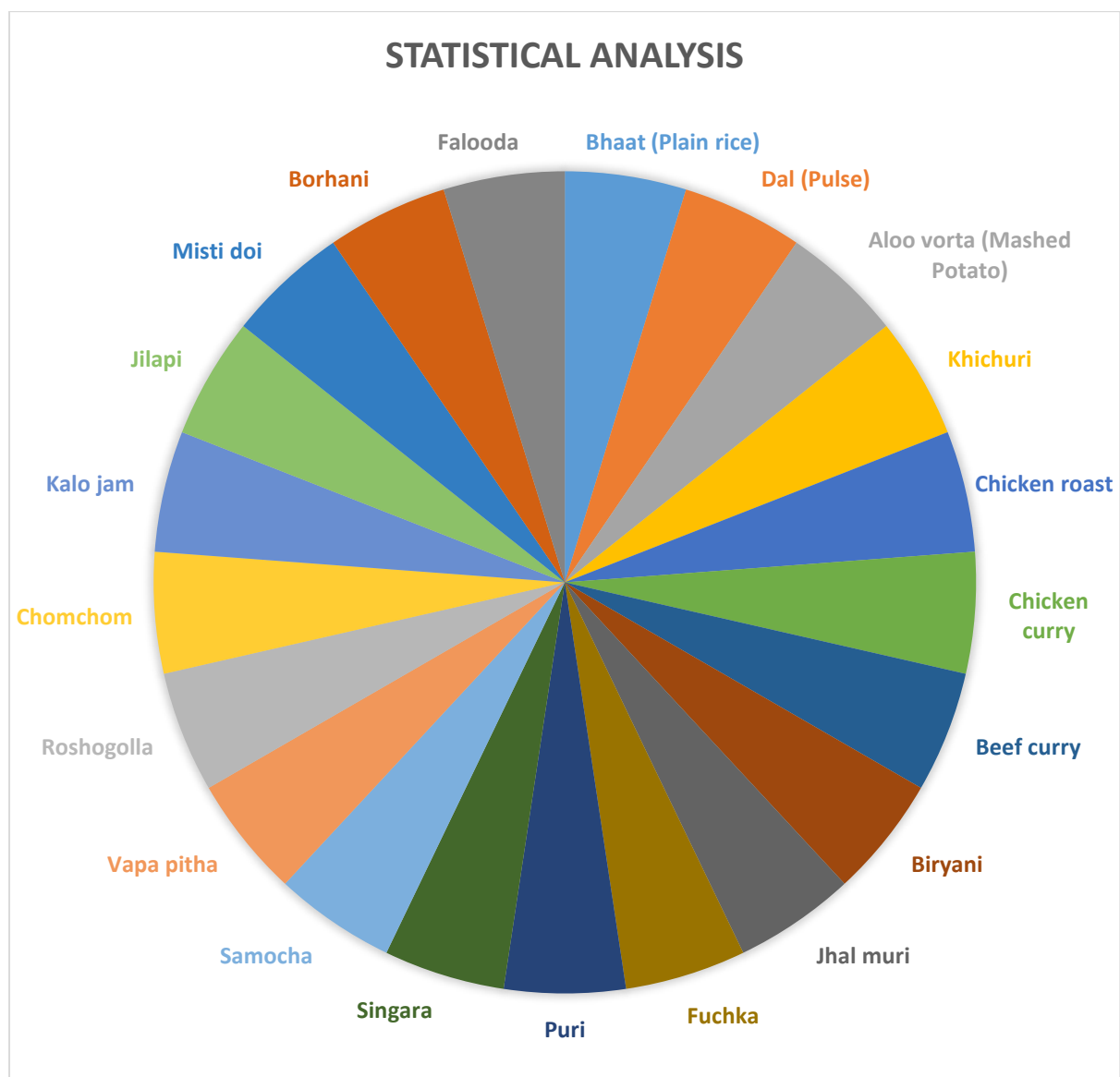


Figure 3.3: Statistical Analysis of the Dataset

3.4 Proposed Methodology

The detailed steps for the machine learning approach are as following:

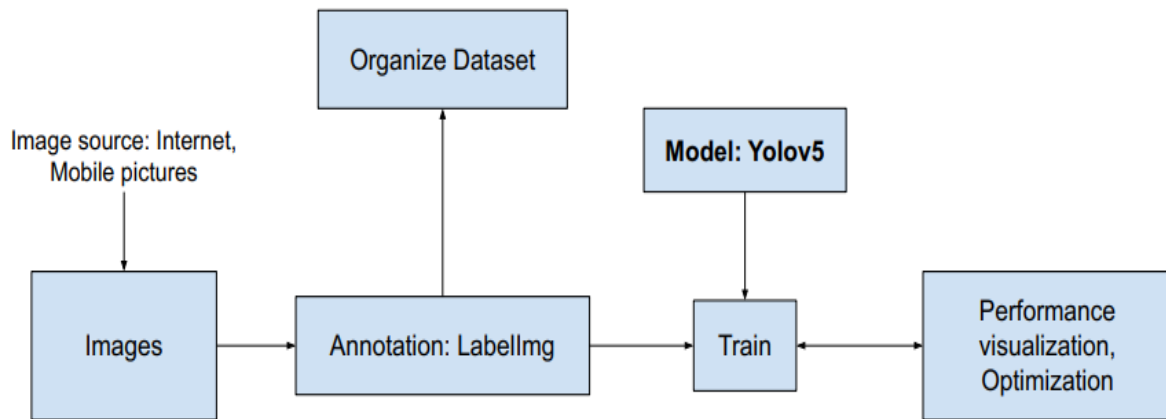


Figure 3.4: Machine Learning Approach

Image files are the source in the machine learning process. The picture data is then cleaned and labelled before being visualized. I produced the organized dataset from pre-processing. Following pre-processing, the model is trained using Yolov5 on an ordered dataset. During model training, the performance score may be visualized and optimized.

Yolo (“You Only Look Once”) is an effective object detection algorithm. There are five versions of this model including Yolo, Yolov2, Yolov3, Yolov4 and Yolov5. For training my dataset, we used Yolov5. Yolov5 is a state of the art object detection algorithm which is used for training image datasets (Świeżewski, 2020). There are four versions of Yolov5: Yolov5s, Yolov5m, Yolov5l and Yolov5x.

3.4.1 Architecture of Yolov5

Yolo, being a real-time object identification framework, is quite capable of properly recognizing foods. It is a simple object detector. That is, it skips the region purpose stage and just predicts across a small number of bounding boxes. RCNN and R-FCN, on the other hand, are incapable of doing so. YOLOv5 is divided into three key sections.

Model Backbone: It extracts key characteristics from the provided input picture. A convolution neural network accumulates and produces visual features at various granularities. Backbone's job is to extract the relevant information from photos. Backbones of many types exist, including Darknet, CSPNet, and others (Solawetz, 2020).

Model Neck: This portion produces feature pyramids, which help in the identification of the same item in multiple sizes and scales. A stack of layers used to mix and blend picture characteristics before passing them on to prediction. Neck's job is to generate features pyramid. Which is capable of dealing with multiscale appropriately. Many modules are recommended for the feature pyramid. As an example, consider the Path Aggregation Network (Solawetz, 2020).

Model Head: Consumes neck features and performs box and class prediction procedures. Anchor boxes are applied to features to provide final output vectors with class probabilities, object ness scores, and bounding boxes. Well, one of the most critical characteristics we can modify for optimal performance on our dataset is anchor boxes. In fact, if anchor boxes are not properly set, our neural network will never know that certain small, huge, or irregular items exist, let alone detect them (Solawetz, 2020).

The approach "just looks once" at the picture in the sense that it provides predictions after just single forward propagation pass through the neural network. It then provides identified items after non-max suppression (which ensures that the object detection algorithm only identifies each object once). The diagram of Object Detection Process is given below (Garg, 2021):

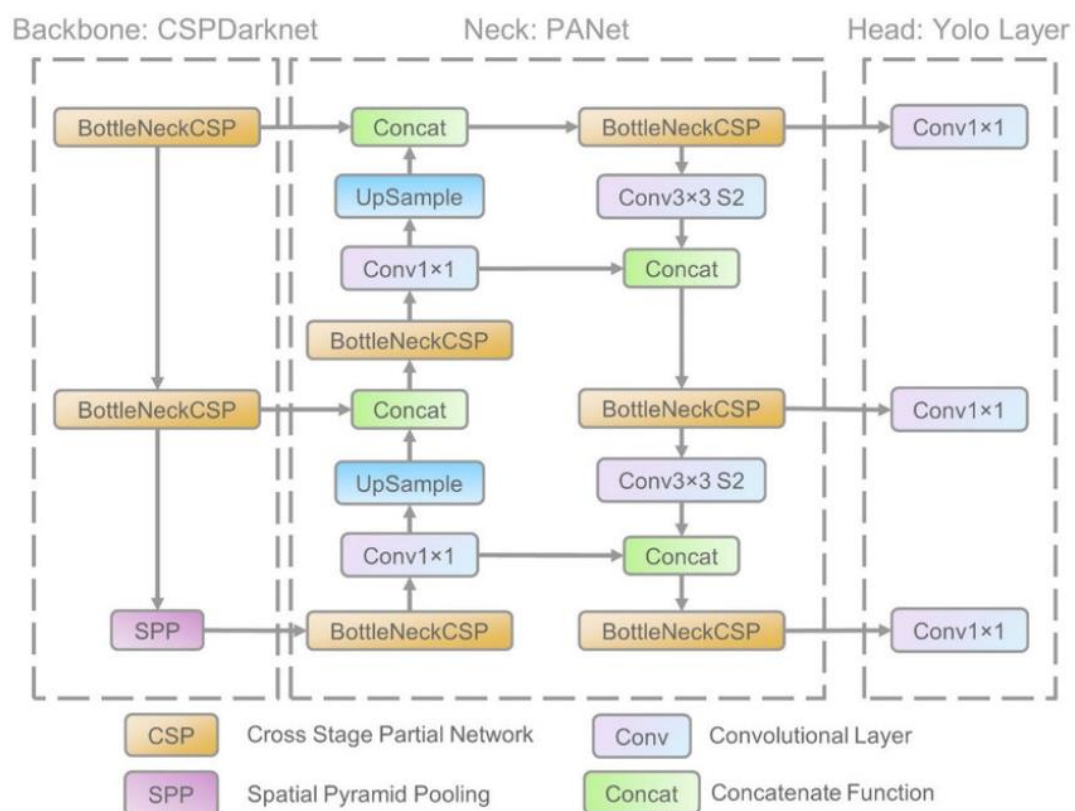


Figure 3.4.1: YOLOv5 Architecture

3.5 Implementation Requirements

The requirements for implementing my project are as follows:

1. Dataset.
2. Training model.
3. IDE.
4. GPU.

Chapter 4

Experimental Results and Discussion

4.1 Experimental Setup

The experiment was carried out on “Google Colaboratory” which is commonly known as “Google Colab”. It is an IDE which lets us execute codes online. It provides a GPU (Graphics Processing Unit) which helps to train custom datasets online. For my experiment I had uploaded my dataset on a website named Roboflow and generated a key to fetch the dataset from Roboflow and implement in Google Colab. I split the dataset into three parts.

1. Training set – 80%.
2. Validation set – 10%.
3. Testing set – 10%.

After splitting the dataset, I trained the dataset for 100 epochs. The images were resized to 512 * 512 pixels for training the model.

4.2 Experimental Results and Analysis

There are four YOLOv5 models. I have trained my dataset on YOLOv5s, YOLOv5m and YOLOv5l. After training the models several metrics and graphs were generated which lets us to evaluate the performance of the training model. Here, the performance on the three models are given discussed below:

4.2.1 YOLOv5s

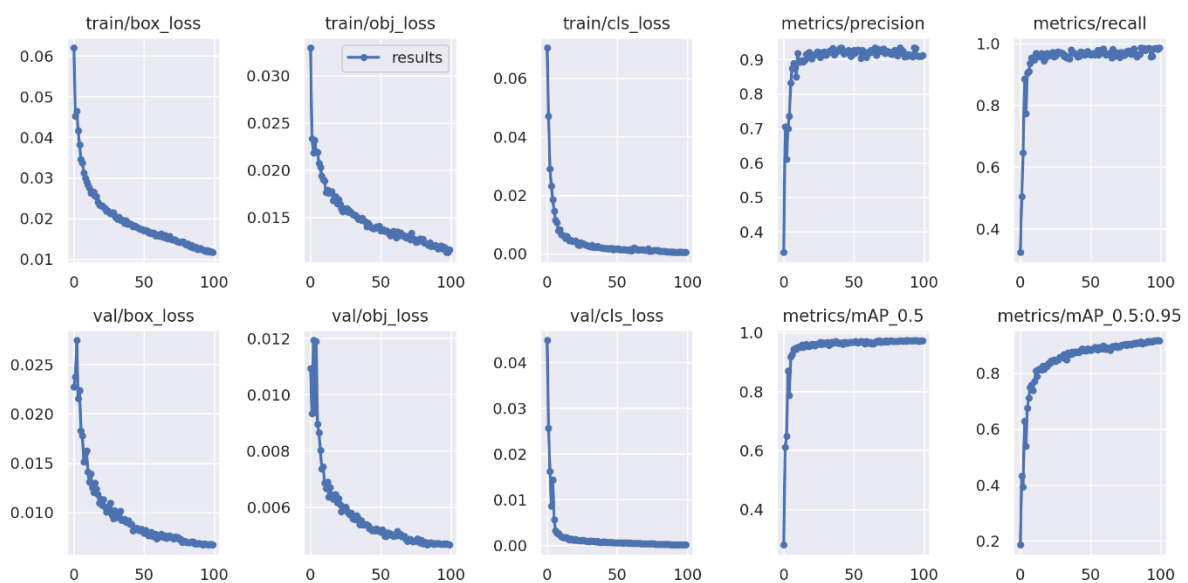


Figure 4.2.1: YOLOv5s Evaluation metrics

In Figure 4.2.1 we can see that four metrics were generated during training the model. In mAP_0.5 we can see that the model has a Mean Average Precision of about 97.3% while working on 50% of the dataset.

In mAP_0.5:0.95 we can see that the model has a Mean Average Precision of about 92.1% while working on 50% - 95% of the dataset.

The precision metric shows a value of 90.3% at the end of the training.

The recall metric has a performance value of 98.7%.

In Figure 4.2.1 we can see that during the training and validation of the dataset, the box_loss, cls_loss and obj_loss graphs are going down as the number of epochs increased.

4.2.2 YOLOv5m

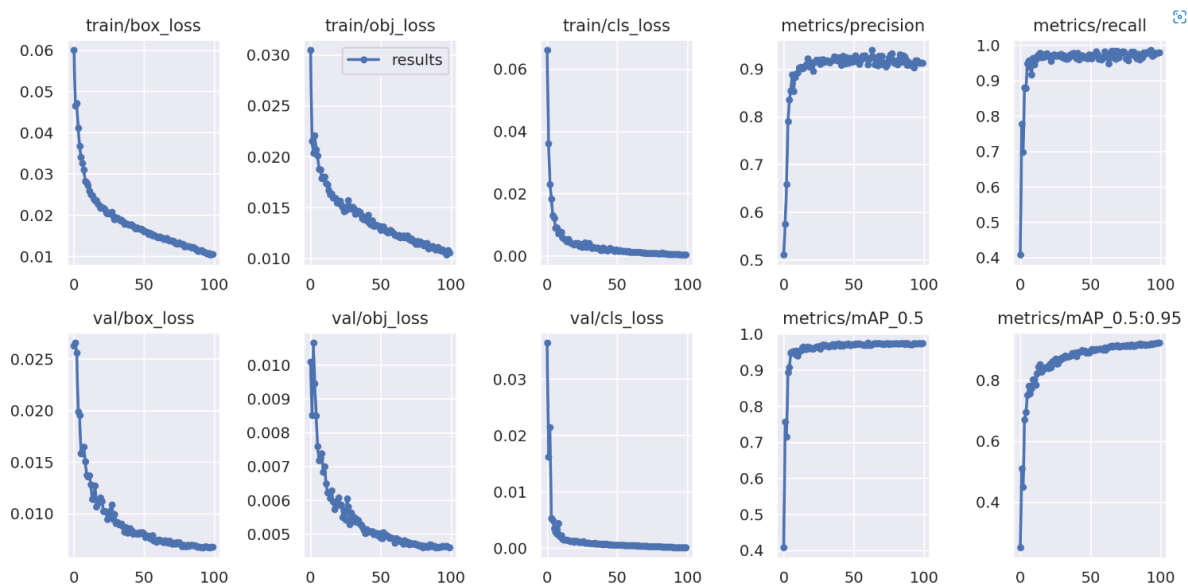


Figure 4.2.2: YOLOv5m Evaluation metrics

In Figure 4.2.2 we can see that four metrics were generated during training the model. In mAP_0.5 we can see that the model has a Mean Average Precision of about 97.4% while working on 50% of the dataset.

In mAP_0.5:0.95 we can see that the model has a Mean Average Precision of about 92.2% while working on 50% - 95% of the dataset.

The precision metric shows a value of 91.3% at the end of the training.

The recall metric has a performance value of 97.9%.

In Figure 4.2.2 we can see that during the training and validation of the dataset, the box_loss, cls_loss and obj_loss graphs are going down as the number of epochs increased.

4.2.3 Yolov5l

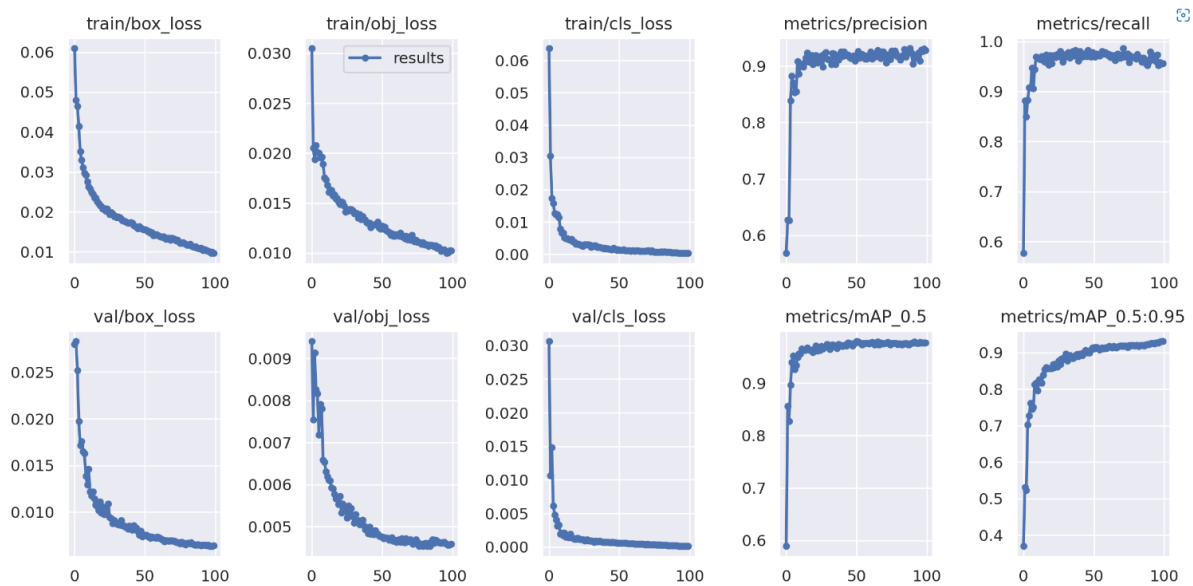


Figure 4.2.3: Yolov5l Evaluation metrics

In Figure 4.2.3 we can see that four metrics were generated during training the model. In mAP_0.5 we can see that the model has a Mean Average Precision of about 97.7% while working on 50% of the dataset.

In mAP_0.5:0.95 we can see that the model has a Mean Average Precision of about 93.2% while working on 50% - 95% of the dataset.

The precision metric shows a value of 92.8% at the end of the training.

The recall metric has a performance value of 95.7%.

In Figure 4.2.3 we can see that during the training and validation of the dataset, the box_loss, cls_loss and obj_loss graphs are going down as the number of epochs increased.

4.2.4 Comparative Analysis of the training models

In a table below, the comparison among the three training models is visualized:

Table 4.2.4: Comparative Analysis of the training models

Metrics	Yolov5s	Yolov5m	Yolov5l
Precision	93 %	92.5 %	93 %
Recall	95.5 %	96.1 %	95.8 %
F1 Score	94 %	94 %	94 %
mAP	94.15 %	94.1 %	95.3 %

From the given table we can see that the Yolov5l model has the best Precision and mean average precision score which are respectively 93% and 95.8 % with a recall value of 95.8 %. The Yolov5m has gained the best recall value which is 96.1 %. The Yolov5s has Precision score of 93% and mean average precision of 94.15%. If we look at the numbers, the Yolov5l model has the overall best performance on my dataset.

4.2.5 Precision-Recall curve analysis

The precision-recall curve represents the tradeoff between precision and recall for various threshold values. A high area under the P-R curve indicates both strong recall and high accuracy, while high precision corresponding to a low false positive rate and high recall corresponding to a low false negative rate. High scores for both indicate that the classifier is producing accurate results (high precision) and producing the majority of all positive outcomes (high recall) (Precision-Recall, n.d.).

After the training of the dataset on three models, different Precision-Recall curves were generated on the testing data. The Precision-Recall curves for the three training models are given below.

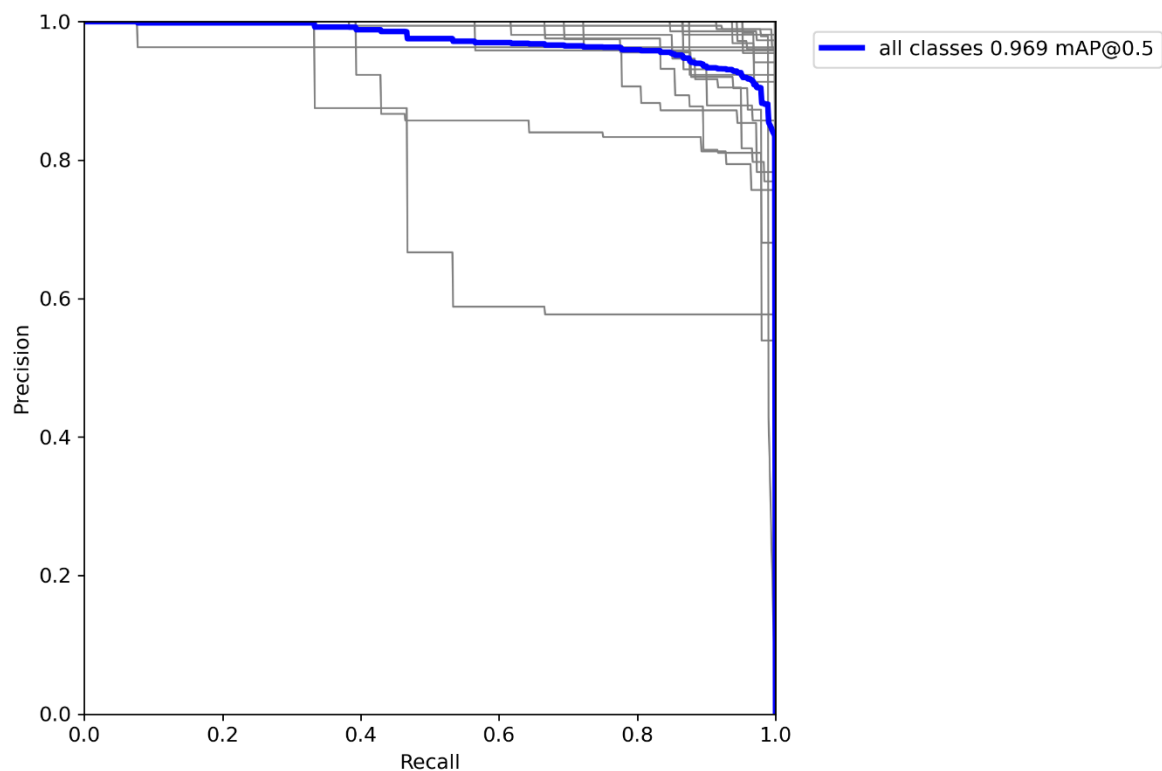


Fig 4.2.5 (1): Precision-Recall curve for Yolov5s

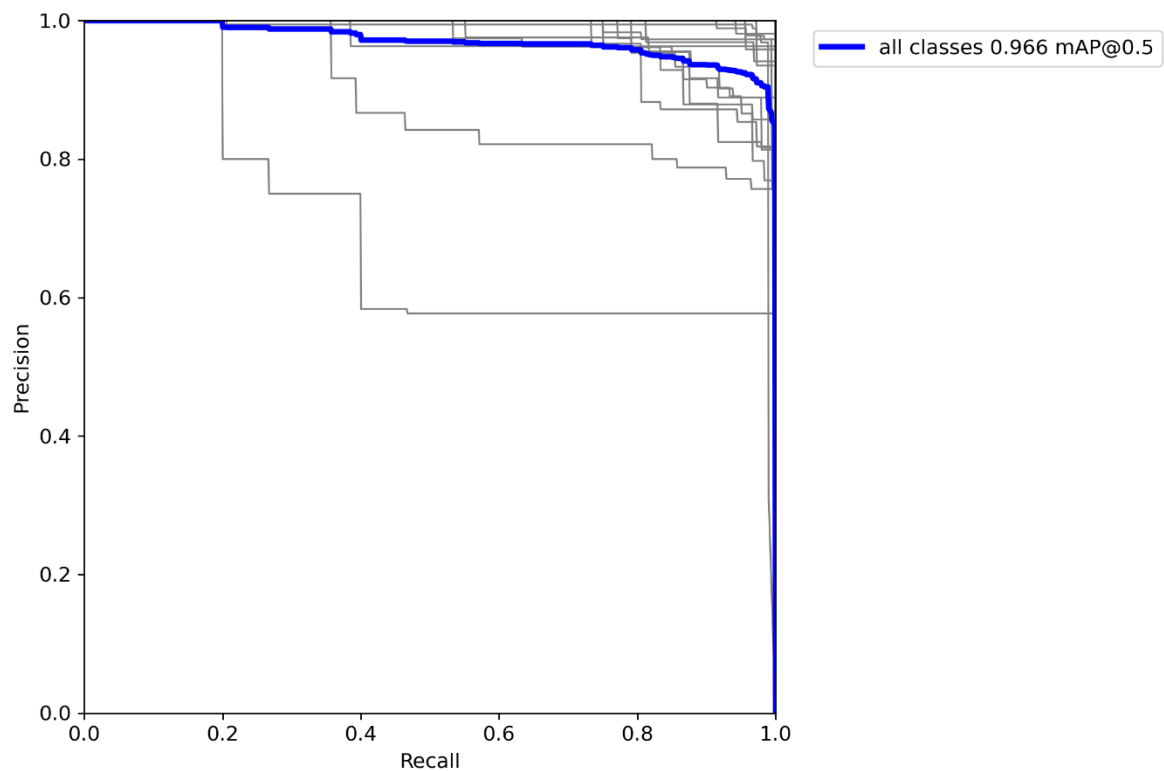


Fig 4.2.5 (2): Precision-Recall curve for YOLOv5m

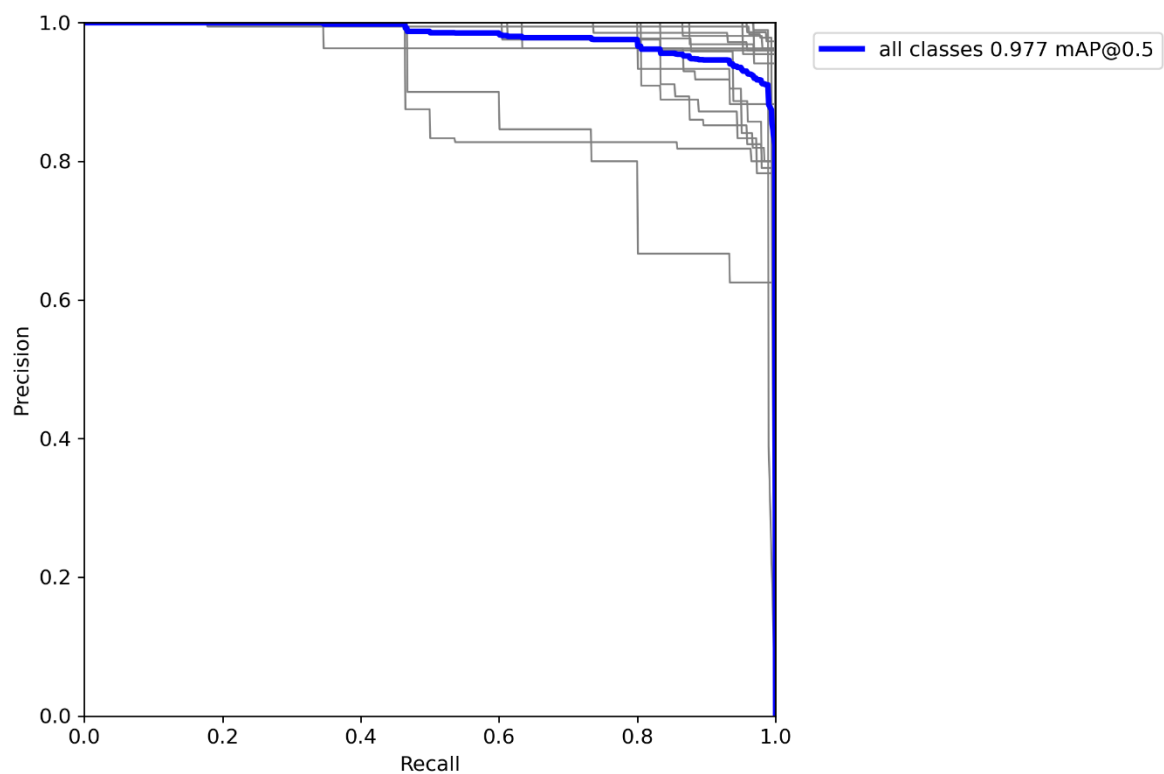


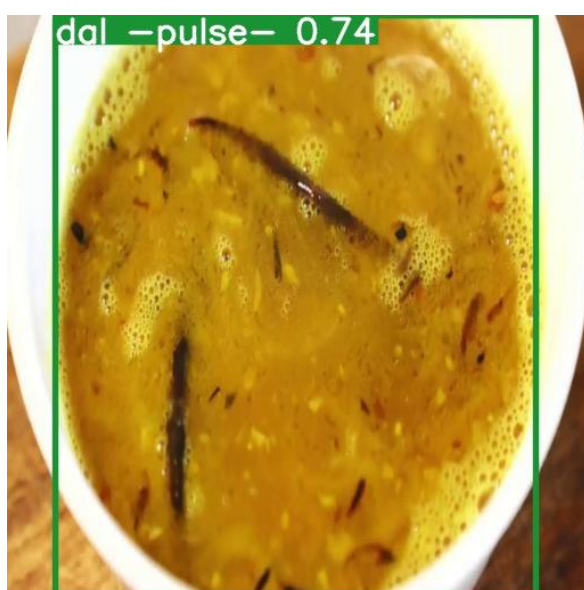
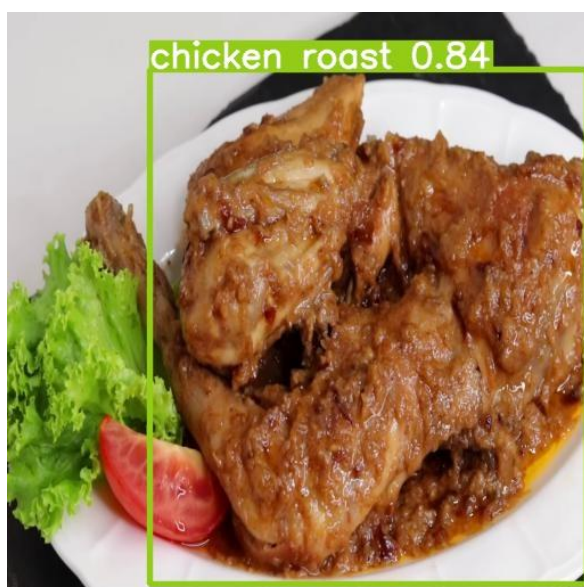
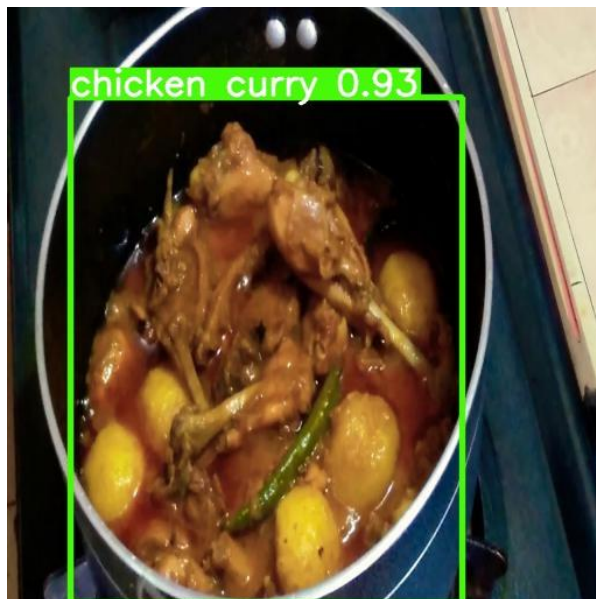
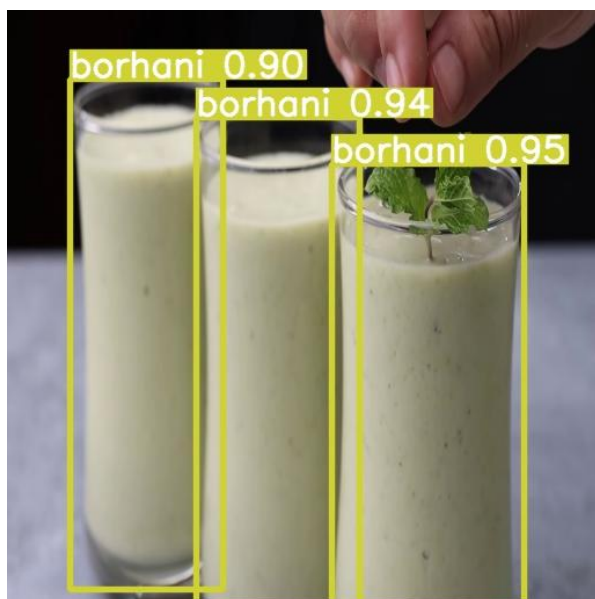
Fig 4.2.5 (3): Precision-Recall curve for YOLOv5l

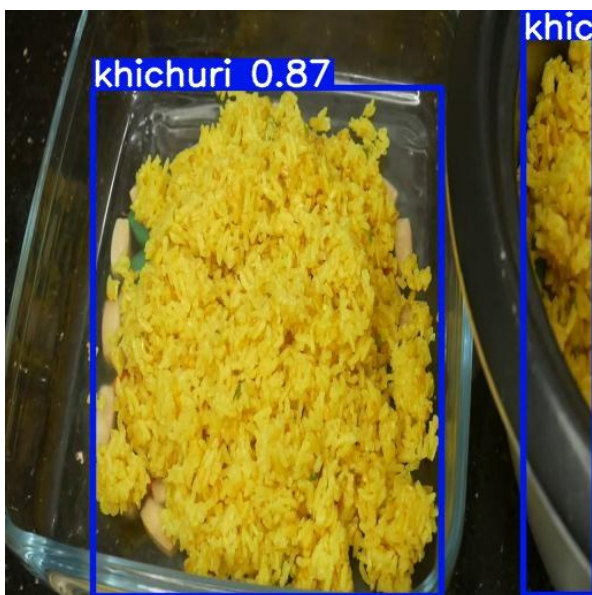
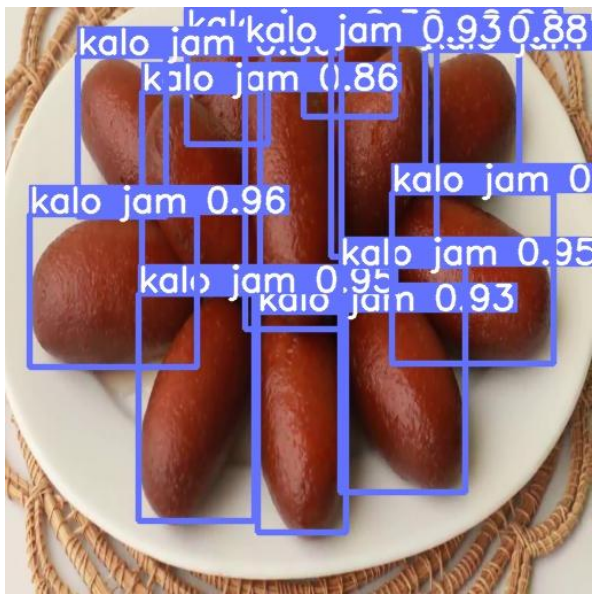
From the Precision-Recall curves above, it is noticed that the Yolov5l model has the best performance on my dataset with a score of 97.7%.

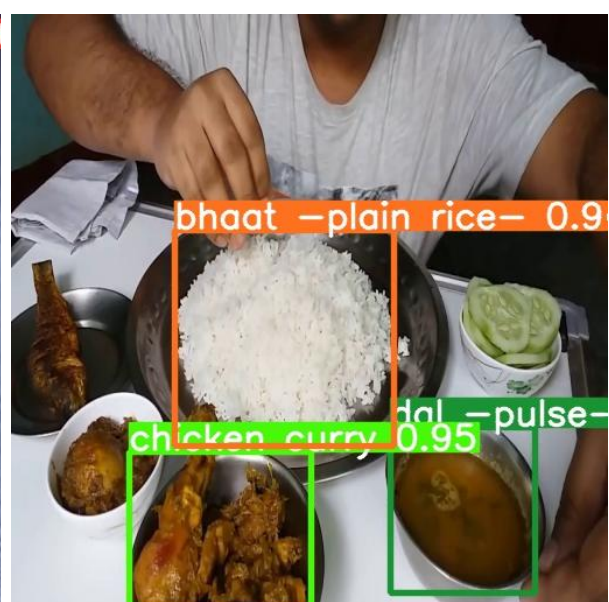
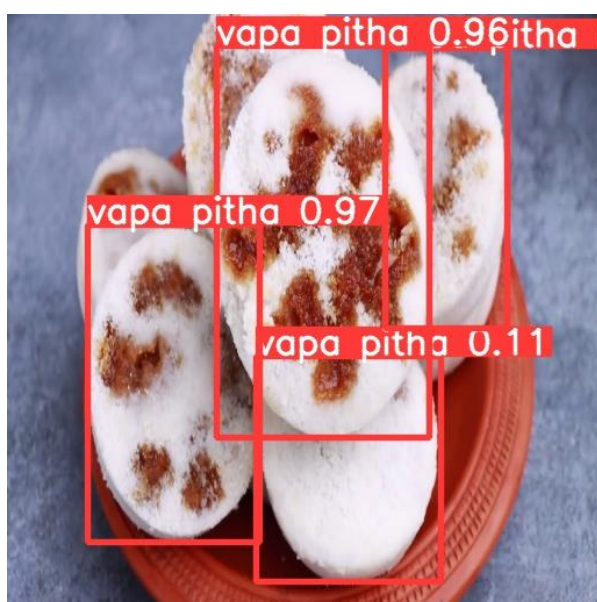
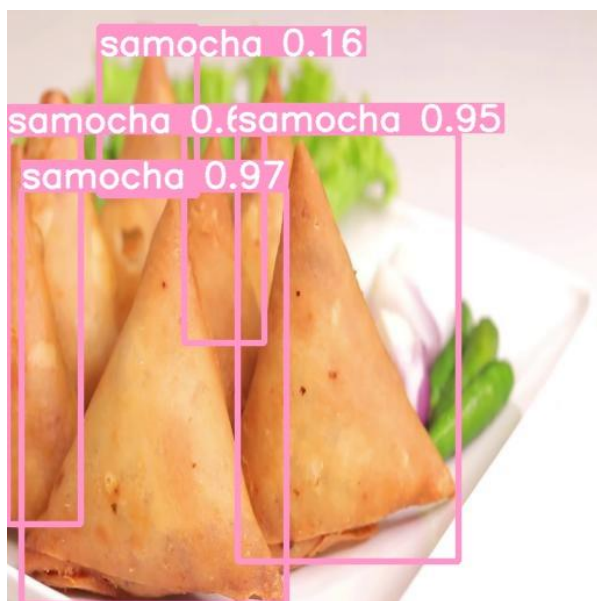
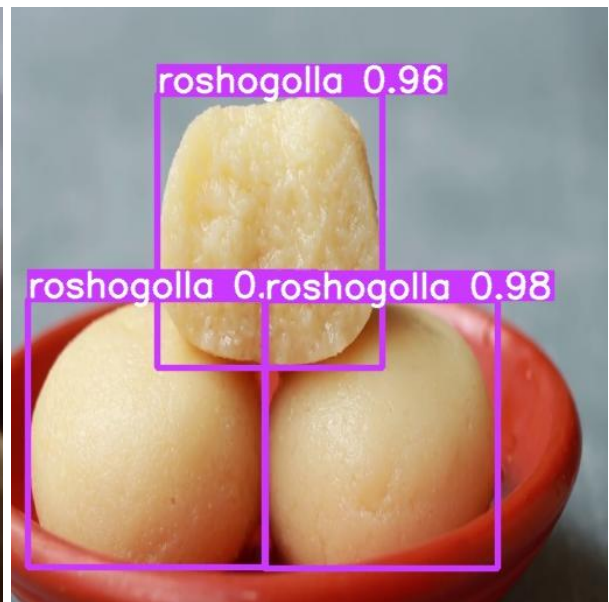
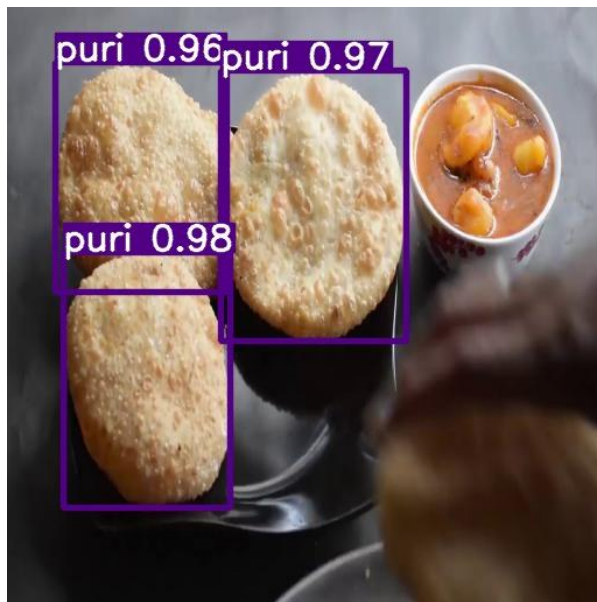
4.3 Discussion

After completing the training of the model, the performance was evaluated. In the confusion matrix shows, it has satisfactory performance in predicting the food classes. The model performance was evaluated throughout 100 epochs and the performance improved as the epochs increased. The outputs on the test set are given below:









Chapter 5

Impact on Society, Environment and Sustainability

5.1 Impact on Society

Machine learning (ML) is an Artificial Intelligence (AI) technology that allows the system to automatically learn and improve through experience rather than explicit programming. This is feasible because enormous amounts of data are now available, allowing robots to be learned rather than programmed. It is seen as a huge technical advance capable of analysing large amounts of data. Almost everything that comes out of the technological industry these days appears to have an aspect of artificial intelligence or machine learning (ML). ML today is the application of some parts of human talents, but not the complete potential for human intelligence. By interacting with intelligent software, ML empowers individuals to do more. It's akin to giving technology a more human face (Singh, 2020).

Machine learning is reshaping the world through transforming industries such as healthcare, education, transportation, food, entertainment, and various assembly lines, among others. It will have an influence on practically every element of people's life, including homes, automobiles, shopping, food ordering, and so on. Technologies such as the Internet of Things (IoT) and cloud computing are all increasing the use of ML to make items and devices "smart" for themselves. Companies attempting to use big data for customer happiness may find ML useful (Singh, 2020).

This project will have impacts on the society. If the model can be implemented into a mobile application, it can be useful. The app can be used for children to get to know Bangladeshi dishes as well as for the tourists who visit this country. They can get recognise Bangladeshi foods they are unfamiliar with.

5.2 Impact on Environment

Machine learning is at the heart of artificial intelligence. It enables AI to develop and adapt to new changes. This type of learning uses algorithms to teach itself and rapidly achieve self-sufficiency. Machine learning, with its huge potential and capacity, may now assist specialists in environmental preservation and wildlife conservation. Natural catastrophes, global warming, and human impact all have different effects on the ecosystem. Storms may contaminate water infrastructure, and animals are constantly threatened by the climate catastrophe and human hunting. Though machine learning cannot prevent every of these incidents, it can lessen their frequency and the harm they cause (Folk, 2021).

5.3 Ethical Aspects

The ethical aspects regarded with Machine Learning are non-violation of privacy, doing no harm to anyone, confidentiality, assessing relevant components only etc. ("Ethical considerations", 2022). During the development of this project, the factors were considered and followed carefully.

Chapter 6

Summary, Conclusion, Recommendation and Implication for Future Research

6.1 Summary of the Study

The name of the project is “**Bangladeshi Food Detection Using Yolov5**”. The project was chosen to categorize Bangladeshi foods and to create a dataset for a wide range of Bangladeshi dishes. The project can be divided into four sections:

1. Background Study.
2. Data Collection.
3. Data Processing.
4. Model Training.

Before starting to work in this project, I did some research on object detection primarily on food detection. After gathering information, I started collecting data and processed them. After the dataset got ready, I trained the Yolov5 model for my custom dataset and evaluated the results. During making the dataset, I have categorised 21 categories of images and processed 240 images for each category, which makes the dataset size 5040.

6.2 Conclusion

The research started with a vision of implementing Machine Learning technique in a different area which was not much explored. There have been works on food recognition around the world. But there have been very less works done on the field of Bangladeshi foods. So exploring this category is very important. In this project, I have categorised 21 different types of Bangladeshi dishes, but there are many more dishes which need to be addressed. Datasets are one of the most valuable assets in current world. This dataset will be available online and can be used by people for different research studies and development purposes.

6.3 Implication for Further Study

Object detection is one of the most popular and diverse field of study. The dataset that I have created can be used for different purposes. This dataset can be used for dietary assessment, calorie estimation, automatic food price estimation etc. There are further plans about the training model. I will be working on an Android Application where this trained model will be implicated to recognise Bangladeshi foods.

References

1. Giovany, S., Putra, A., Hariawan, A., & Wulandhari, L. (2017). *Machine Learning and SIFT Approach for Indonesian Food Image Recognition*.
2. Suma, S., Bharathi, M., Harikumar, G., & Deepak, B. (2020). *Food Recognition And Calorie Estimation Using Image Processing*.
3. Uddin, M. A., Miraj, A. A., Sarma, S. M., Das, A., & Gani, M. (2021). *Traditional Bengali Food Classification Using Convolutional Neural Network*.
4. Jeny, A., Junayed, M., Ahmed, I., Habib, M., & Rahman, M. (2019). *FoNet - Local Food Recognition Using Deep Residual Neural Networks*.
5. Bonis, D. M., Amato, G., Falchi, F., Gennaro, C., & Manghi, P. (n.d.), *Deep Learning Techniques for Visual Food Recognition on a Mobile App*.
6. Priyanka, N., & Geetha, V. (2022), *Food Recognition Using Deep Convolutional Neural Network*.
7. Świeżewski, J. (2020). *YOLO Algorithm and YOLO Object Detection*. Appsilon. Retrieved 20 May 2022, from <https://appsilon.com/object-detection-yolo-algorithm/>.
8. Solawetz, J. (2020). *YOLOv5 New Version - Improvements And Evaluation*. Roboflow Blog. Retrieved 20 May 2022, from <https://blog.roboflow.com/yolov5-improvements-and-evaluation/>.
9. Garg, A. (2021). *How to Use Yolo v5 Object Detection Algorithm for Custom Object Detection*. Analytics Vidhya. Retrieved 20 May 2022, from <https://www.analyticsvidhya.com/blog/2021/12/how-to-use-yolo-v5-object-detection-algorithm-for-custom-object-detection-an-example-use-case/>.
10. Singh, V. (2020). *How Machine Learning Is Changing the World*. Data Science Central. Retrieved 20 May 2022, from <https://www.datasciencecentral.com/how-machine-learning-is-changing-the-world/>.
11. Folk, E. (2021). *How Machine Learning Can Help in Environment Protection?*. Techie Loops. Retrieved 20 May 2022, from <https://techieloops.com/how-machine-learning-can-help-in-environmental-protection/>.
12. *Ethical considerations*. My-peer toolkit. (2022). Retrieved 23 May 2022, from <https://mypeer.org.au/monitoring-evaluation/ethical-considerations/>.
13. *Precision-Recall*. scikit-learn. Retrieved 8 June 2022, from https://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html#:~:text=The%20precision%2Drecall%20curve%20shows,a%20low%20false%20negative%20rate.