# mDiet: A DIETARY MANAGEMENT SYSTEM OF DIABETES IN PREGNANCY USING MACHINE LEARNING

**BY**

**TAMIM AHAMED**
**ID: 201-15-13976**
**AND**

**SRABONI AKTER**
**ID: 201-15-13746**

## FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

## Supervised By

**Dr. Md. Taimur Ahad**
Associate Professor and Associate Head
Department of Computer Science and Engineering
Daffodil International University

## Co-Supervised By

**Ms. Sharun Akter Khushbu**
Sr. Lecturer
Department of Computer Science and Engineering
Daffodil International University

## DAFFODIL INTERNATIONAL UNIVERSITY

**DHAKA, BANGLADESH**

**July 2024**

## APPROVAL

This Project titled **"mDiet: A Dietary Management System Of Diabetes In Pregnancy Using Machine Learning"**, submitted by **Tamim Ahamed** and **Sraboni Akter** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **15 July 2024.**

## BOARD OF EXAMINERS

**Dr. Md. Ismail Jabiullah (MIJ)** — **Board Chairman**
**Professor**
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

**Mr. Md Assaduzzaman (MA)** — **Internal Examiner 1**
**Sr. Lecturer**
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

**Ms. Sharun Akter Khushbu (SAK)** — **Internal Examiner 2**
**Sr. Lecturer**
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

**Dr. Ahmed Wasif Reza (DWR)** — **External Examiner**
**Professor**
Department of Computer Science and Engineering
East West University

## DECLARATION

We hereby declare that this project has been done by us under the supervision of **Dr. Md. Taimur Ahad, Associate Professor and Associate Head, Department of Computer Science and Engineering, Daffodil International University**. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

**Supervised by:**

Taimur Ahad

**Dr. Md. Taimur Ahad**
Associate Professor and Associate Head
Department of CSE
Daffodil International University

**Co-Supervised by:**

Khushbu

**Ms. Sharun Akter Khushbu**
Sr. Lecturer
Department of CSE
Daffodil International University

**Submitted by:**

Tamim

**(Tamim Ahamed)**
ID: 201-15-13976
Department of CSE
Daffodil International University

Sraboni Akter

**(Sraboni Akter)**
ID: 201-15-13746
Department of CSE
Daffodil International University

② ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty for His divine blessing making it possible for us to complete the final year project/internship successfully.

We are grateful and wish our profound indebtedness to **Dr. Md. Taimur Ahad**, **Associate Professor and Associate Head**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Machine Learning*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartiest gratitude to the **Dr. Sheak Rashed Haider Noori, Professor & Head, Department of CSE**, for his kind help in finishing our project and also to other faculty members and the staff of the Department of CSE, Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

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

# ABSTRACT

This paper proposes a machine-learning approach that leverages deep-learning techniques specifically using the YOLOv8 model for real-time food item detection from images and recommendations, helping pregnant women with diabetes manage their diets effectively. As a dataset, 5,016 raw images with 18 different food classes were collected. All the images were captured from different sources with various backgrounds and lighting conditions. Several preprocessing techniques were also used such as annotation, augmentation, etc. to prepare the dataset. By using online annotating tools several bounding boxes were drawn around the food items. To make the balance between food classes and increase the dataset size augmentation techniques were also performed. With the prepared dataset two versions (YOLOv5 and YOLOv8) of the YOLO family are used. Both of the models performed well and were able to detect objects on food images accurately. But comparatively, the YOLOv8 model achieved the highest mean Average Precision (mAP). The model achieved a 0.882 mAP50 score and a 0.709 mAP50-90 score which indicates that our model is capable of detecting food items from images pretty well. As our proposed method has a recommendation system this model was later used on a mobile application called mDiet. In our mDiet application, a recommendation system has been integrated that takes user health and pregnancy-related information as input, which further relates to detected food class, because the model is also integrated to detect food items from images by capturing food images using a mobile camera or selecting images from their device. By providing their information and food images pregnant women could benefit from the mDiet application in diet management and get useful personalized recommendations. Overall using the YOLOv8 model and the mDiet application we got our expected result.

## TABLE OF CONTENTS

## LIST OF FIGURES

# LIST OF TABLES

# CHAPTER 1
# INTRODUCTION

## 1.1 Overview

Ensuring that the mother and developing fetus are consuming the correct amount of food during pregnancy is crucial for maternal health and nutrition. When a pregnant woman is diagnosed with gestational diabetes mellitus, which is characterized by high glucose levels throughout pregnancy, proper nutrition becomes even more challenging. Making sensible, health-conscious lifestyle choices can lower blood glucose levels, which lowers risk and promotes good health. To control gestational diabetes mellitus and stay away from excess calories, a balanced diet is crucial. The formula has healthy benefits and it is moderate in low energy and high in vitamins, minerals, and fiber. It is recommended that pregnant women choose diets low in fat and carbohydrates to meet their demands while maintaining weight increase. Maintaining excellent health and controlling blood sugar can also be aided by regular exercise and a balanced diet.

In this work, we use the YOLOv8 model to build accurate food recognition and evaluate eating patterns in pregnant women with gestational diabetes mellitus. The system can detect different food items from photo scans and provides real-time nutritional suggestions. For each food, the deep learning models are fed the nutritional value and visual characteristics Nutritional management and diabetes treatment have long faced the difficulty of identifying food items and evaluating their dietary impact. Modern deep learning methods and nutritional science are closely connected in our efforts to solve these issues.

The project aims to encourage pregnant mothers to make knowledgeable and healthful dietary choices throughout pregnancy by combining modern technologies with nutritional assessment. A food recommendation system designed for pregnant women is an attempt to address the issues raised by this study. The device uses image recognition technology to identify scanned food images to provide real-time dietary recommendations.

## 1.2 Background and Present State

The importance of proper nutrition during pregnancy cannot be overstated, as it plays a crucial role in the health and development of both the mother and the fetus. Managing food intake to meet nutritional demands and maintain blood sugar levels presents special problems for pregnant women, especially those with gestational diabetes. Manually documenting food consumption is a common component of traditional dietary control techniques, which can lead to mistakes including estimating portion sizes incorrectly and providing insufficient nutritional information.

Technological developments in the last several years have created new opportunities to enhance nutritional control techniques. For example, technology such as computer vision and deep learning have shown great potential in several applications, such as dietary evaluation and food recognition. Modern object identification models like YOLO (You Only Look Once) have shown amazing results in the real-time processing of images. The latest iteration, YOLOv8, offers enhanced accuracy and speed, making it an ideal tool for developing an automated food recommendation system.

There is still not enough use of these technological advancements to help pregnant women manage their diets, despite the advances in technology. Our goal is to close this gap by creating a comprehensive meal recommendation system that uses YOLOv8 to classify and recognize foods in real time. Expectant women will be able to make accurate and timely food decisions with the support of the system's individualized dietary suggestions.

### 1.3 Problem Statement

Accumulating evidence has suggested that dietary intake before and during pregnancy is associated with the risk of GDM [1]. The healthfulness and nutritional value of different foods provide obstacles for pregnant women when making dietary decisions. A frequent pregnancy problem known as diabetes occurs when pregnant women without diabetes experience glucose intolerance. Depending on the diagnostic criteria and screening methods employed, estimates of its prevalence around the world range from less than 1 to 28%, although they have increased by more than 30% over the previous 20 years.

For instance, several cross-sectional or retrospective studies have shown that high intakes of saturated fat, cholesterol, or eggs are associated with an increased risk of diabetics and polyunsaturated fat intake appears to have a protective effect Pregnant women are unclear and confused about their nutritional choices since existing dietary guidance systems lack personalization and real-time feedback [2].To further highlight the necessity for individualized dietary recommendations to reduce pregnancy problems, consider the alarming increase in the prevalence of gestational diabetes mellitus.

To address these problems, a technology-based solution that offers customized nutritional advice to reduce pregnancy-related difficulties is desperately needed. Such a system can provide accurate and customized advice, thereby improving nutritional management for pregnant mothers and ultimately leading to better health outcomes for mothers and children. It does this by utilizing the latest algorithms for deep learning algorithms and real-time picture recognition technology.

## 1.4 Objectives

Pregnant women can use this system to receive personalized dietary suggestions quickly, based on their requirements and preferences. Its objective is to accurately detect and classify 18 different food types. Although it is our goal to design this new system securely, we also want to make sure that pregnant mothers can use it in the real world to make comfortable and trained food choices.

Our work aims to make the difficult chore of tracking nutrients and making decisions during pregnancy easier by providing immediate feedback on the acceptability and safety of food selections. After all, we want to minimize the dangers of poor nutrition control and promote the best possible health results for the mother and the baby. Automated evaluation solutions like YOLOv8 apply cutting-edge deep learning and computer vision technologies and provide fast assistance. These advanced algorithms assess the meal's nutritional value, provide tailored dietary suggestions, and recognize food items from scanned photos with speed and accuracy. The program seeks to automate the nutritional control process for expectant moms with gestational diabetes by utilizing YOLOv8 for real-time meal detection.

Our method uses computer vision to evaluate scanned food photographs and provides real-time dietary suggestions that exactly match the nutritional needs of expectant moms. We hope that by using this creative strategy, pregnant mothers will be more capable of making informed food choices during their pregnancy, improving both the mother's and the baby's health. Pregnant women typically have to manually keep track of their food intake, assess nutritional data, and modify their diets to comply with recommended intake levels. This approach is not only difficult but also prone to mistakes like inaccurate or insufficient portion calculation.

## 1.5 Scope and Limitations

Food identification and dietary analysis using YOLO models have been used to treat gestational diabetes mellitus in pregnant women. However, the healthcare sector is becoming more and more interested in utilizing machine learning models. The effectiveness of YOLO and related models for tasks such as object detection has been demonstrated in numerous studies across various domains; however, there is a dearth of research on tailoring YOLO for precise food item classification and analysis in real-world scenarios, particularly for dietary management applications. Research has demonstrated the potential advantages of models like YOLO for broad photo recognition tasks. However, there is a lack of study on the use of these models to precisely identify specific items ingested by pregnant women and produce dietary recommendations.

Furthermore, while the use of technology in the management of persistent diseases like diabetes has advanced significantly, it is still relatively new when it comes to the control of food during pregnancy. To treat and prevent gestational diabetes mellitus, which benefits the health of both the mother and the baby, individualized dietary suggestions that include the particular dietary requirements and preferences of expectant mothers can be exceptionally beneficial.

The system can identify and categorize food items into 18 different classifications. Using their smartphone, pregnant women can take pictures of their food, and the app will recognize the food items and advise whether it's safe to eat or not based on their dietary requirements and preferences. With this quick input, pregnant mothers may make informed meal decisions in real-time, supporting appropriate nutritional intake and reducing the chance of health problems associated with pregnancy.

Further, dietary choices and adherence to standards are also influenced by cultural and economic variables. To assist solve these issues and make it simpler for pregnant women from different backgrounds to maintain a healthy diet, an advanced system that can identify different food products and provide culturally acceptable recommendations can be developed.

### 1.6 Report Organization

The report is organized into seven chapters that detail the development and implementation of a real-time food recognition and recommendation system for pregnant women with diabetes. Each chapter focuses on specific aspects of the project:

Chapter 1: This chapter is the Introduction that provides an overview of the project's background, objectives, and significance in improving dietary management for pregnant women with diabetes.

Chapter 2: This is the Literature Review that reviews existing research on computer vision and deep learning techniques, particularly YOLO, in automatic food recognition systems. It discusses the potential and limitations of current technologies in dietary assessments.

Chapter 3: Methodology - It details the methodology used in gathering data, preprocessing images, and training the YOLOv8 model with a custom dataset of 5,016 images across 18 food classes. It also outlines the implementation of the recommendation system and its integration with Firebase.

Chapter 4: Experimental Results - It discusses the experimental setup and provides details including model training configurations and performance evaluation metrics such as precision, recall, and mean Average Precision (mAP). It presents detailed results from training YOLOv5 and YOLOv8 models and their comparative analysis.

Chapter 5: Implementation - this chapter describes the integration of the trained YOLOv8 model into a mobile application (mDiet), converted to Tensor Flow Lite for efficient on-device inference. It includes features such as image capture, real-time food detection, and personalized dietary recommendations.

Chapter 6: Impact Assessment - this chapter discusses the potential impact of the Pregnancy Women's Food Recommendation System on health outcomes, sustainability, and ethical considerations. It discusses the system's social implications and its role in promoting sustainable dietary practices.

Chapter 7: Conclusion and Future Work - It summarizes key findings and conclusions drawn from the project, highlighting achievements, limitations related to dataset specificity and recommendation system parameters, and potential conflicts of interest. It also outlines future directions for expanding the system's capabilities, including adding more food classes and improving the recommendation algorithm.

### 1.7 Summary

Good nutrition is essential for the mother's and the unborn child's health throughout pregnancy, especially for women who have diabetes and must regulate their food to keep their blood glucose levels within normal ranges. Advances in technology such as computer vision and deep learning are changing modern nutritional control procedures by simplifying and reducing errors.

Based on the YOLOv8 model, the project creates an accurate meal recognition system for pregnant mothers with gestational diabetes. Through the use of photos to identify and classify food items and provide real-time nutritional recommendations, the technology assists pregnant mothers in making informed decisions about foods.

With the use of technologies like YOLOv8, this project aims to create a real-time food recommendation system for pregnant mothers, helping them to better control their diets. The system's objectives are to give pregnant women individualized dietary recommendations, simplify nutrition tracking, accurately identify and classify 18 different types of foods, and improve pregnant mothers' health.

The implementation of such technology in practical situations, especially in pregnancy dietary management, presents lots of challenges. Research on YOLO models for dietary needs, detecting foods, and analysis is limited, and cultural and economic factors influence dietary choices and guidelines, necessitating a culturally appropriate system.

# CHAPTER 2
# LITERATURE REVIEW

## 2.1 Overview

Accurate assessment of dietary intake plays a crucial role in nutrition and health research, facilitating the investigation of the connection between nutrition and human health. However, the human body may become fatally overfed with high-calorie consumption. Many ailments, including obesity, diabetes, breast cancer, and more, are mostly caused by an improper diet, especially in pregnant women. Application-based food identification for nutritional recommendations is a better way to solve this issue. One of the most exciting uses of visual object recognition is identifying food images since it can be used to assess dietary patterns in a medical setting. According to our research, visual characteristics from deep neural networks are used for food recognition in about 66.7% of the examined studies. This technology could be particularly useful for personalized diet management during pregnancy. However, personalized nutrition helps different stages of pregnant women to overcome their problems and lead healthy life. An application with detection technology and a recommendation engine could be a better solution to the problem. More studies on this domain can make the opportunities to overcome the problem.

## 2.2 Related Works

According to Ravelli et al [2], proper dietary intake is important for policy, and health research because it allows the investigation of the relationship between diet and human health. To evaluate the diet, Jiang et al. [5] developed a system for food preparation and analysis. In their method, food items are detected using Faster R-CN, and features are extracted by VGG-16. Supervised and unsupervised deep learning models and to extract two features were to form classes from images of food, based on the state of those items. An answer algorithm makes suitable recommendations based on matching users and their products.

The author of [6] suggested a technique for utilizing local component feature relative index space in connection with quick focus. This method is restricted to 61 diets and is challenging to implement for hand-crafted or composite suggestions. In [7], proposed a suitable food reading system that uses an image interpretation method based on learning to interpret the visual content of food. Enhancing the clarity and interpretability of proposed results is the introduction of a rule-written explanation. This climbed dataset was used for experiments, which showed that the suggested approach outperforms current food recommendation techniques in terms of precision, recall, F1 score, and normalized discounted cumulative gain (NDCG), increasing recommendation quality by 7.35%, 6.70%, 7.32%, and 14.38%, respectively. The YOLO technique was used by Romadhon et al. [8] to construct a system for detecting foods. This system was then expanded upon in Python to estimate the calorie count of prepared food. Average precision, recall, and F1-score values of 0.94, 0.90, and 0.91 were achieved by the system during the model-building process. The results of its tests on hugging faces show an identifiable decrease in performance, with average values for precision, recall, and F1-score being 0.84, 0.32, and 0.41, respectively. Using YOLOv3-tiny, [9] created a system for identifying food materials. The Weber Local Descriptor (WLD) feature was used to evaluate the quality of food ingredients, and a Support Vector Machine (SVM) was used to classify the results.

According to the experimental results, food ingredients could be recognized with an accuracy of 82.02% using YOLOv3-tiny, while food quality could be assessed with an 83.33% accuracy utilizing WLD-SVM [9]. Smart Diet Diary is a smartphone application that tracks food consumption and is designed to help patients and obese people control their diet for better health. Its development and design are highlighted in [10]. Nadeem et al. (2023) used a dataset of 16,000 photos of food products divided into 14 different categories to train a multi-label classifier. In [10], analyze that system's total accuracy was roughly 80.1%, and its average calorie computation was within 10% of the real calorie amount. This was accomplished by using a pre-trained Faster R-CNN model for classification[10]. To accomplish simultaneous multi-object recognition and adaptation, the author of [11] constructed a deep convolutional neural network coupled with YOLO, an advanced detection approach, and was able to achieve approximately 80% mean average precision(mAP). Furthermore, a mobile application with expanded nutritional analysis features was created using the model[11]. Liu et al. (2022) clarify the major sensitivities and references of the proposed model-building architecture, SSD Inception V2 and FastR-CNN Inception ResNet V2. AI learning can be sustained thanks to the human-AI philosophy, which was prompted by the transmission of people in data production and preprocessing. In [12], using a small sample size of training food photos, the system was able to accurately identify three different meal categories. Overall, with an accuracy of 0.87, precision of 0.88, recall of 0.97, and F1-score of 0.93, the model successfully classified single-dish, mixed-dish, and multiple-dish meals. Furthermore, the detector was chosen to be the EfficientNet architecture, which provides fast and highly accurate detection as shown by a mAP of 0.92 and the shortest average inference time of less than 5 seconds[12].

A massive, deep convolutional neural network was trained in [13] to categorize 1.2 million high-resolution photos from the ImageNet LSVRC-2010 competition into 1000 distinct classes. On the test data, the model obtained top-1 and top-5 error rates of

37.5% and 17.0%, respectively, greatly surpassing the prior state-of-the-art. A version of this model was also entered in the ILSVRC-2012 competition, where it came in first place with a top-5 test error rate of 15.3%, whereas the second-place entry had a higher mistake rate of 26.2%. Anthimopoulos et al. [14] introduced an automated food recognition method utilizing the bag-of-features (BoF) paradigm. To construct and evaluate the prototype system, an approximately 5000-photo visual dataset of food was created and classified into 11 types. The optimized system uses a linear support vector machine (SVM) classifier to classify the food images, computes dense local features in the HSV color space using the scale-invariant feature transform (SIFT), and creates a visual dictionary of 10,000 visual words using hierarchical k-means clustering. With a classification accuracy of about 78% on a very challenging image dataset, the system demonstrated the feasibility of the proposed approach.

Chen et al. [15] note that since public health concerns have grown, computer-aided food identification and quantity estimation have attracted increased interest in recent years. Several studies have been proposed, and the identification problem is typically framed as an image categorization or classification challenge. Here, they described how to tackle the problems with feature descriptors in the food recognition task and presented a working method for quantity estimation based on depth data. Food items are represented by combining the SIFT and Local binary pattern feature descriptors, which use sparse coding, with Gabor and color characteristics. [15], trained using a multi-label SVM classifier, and these classifiers are then coupled using multi-class Adaboost methods. The assessment comprises fifty food categories from around the globe, with 100 images from different sources—such as manually taken pictures or online photo albums—in each area. An accuracy of 68.3% is obtained overall for N = 2, 3, and 5, while success at top-N candidates yielded comparable accuracy of 80.6%, 84.8%, and 90.9%, indicating the feasibility of a mobile application. The experimental results show that the proposed methods greatly improve the performance of the original SIFT and LBP feature descriptors.

A collection of food photos from three nearby restaurants was gathered by [16] to assess the technique. Five photographers used a mix of one point-and-shoot camera and six different smartphone types to capture the pictures. Recall rates for fewer than four items will always be below 100% because the Menu-Match dataset has an average of 2.1 food items per plate and a maximum of 4. When five food items from each of the three restaurants' menus were included in the list, the average recall rate was 83%, while when only one restaurant's menu was included, the average recall rate was 92%. This sorting method aids people in choosing the right food items fast. This study shows [17] that food recognition accuracy may be greatly improved by combining features from a

Deep Convolutional Neural Network (DCNN) with conventional hand-crafted visual characteristics, like Fisher Vectors with HoG and color patches. On a food dataset consisting of 100 classes, the studies produced a top-1 accuracy of 72.26% and a top-5 accuracy of 92.00%.[18]. X. Yin et al.[19] in their developed a detection framework combining PENet and YOLOv3 which is (PE-YOLO) and it achieved advanced results, achieving 78.0% in mAP and 53.6 in FPS, respectively, which can adapt to object detection under different low-light conditions. D. Reis et al. [20] achieve in their Real-Time Flying Object Detection with YOLOv8 a mAP50 of 79.2%, mAP50-95 of 68.5%, and an average inference speed of 50 frames per second (fps) on 1080p videos.

Emmanuel et al. [21] stress the significance of keeping an eye on the number of calories ingested from foods daily to reduce the risk of diabetes, heart disease, and obesity. Because food mass varies, current calorie estimators are less accurate. A suggested segmentation approach based on Cauchy, Generalized T-Student, and Wavelet kernel-based Wu-and-Li Index Fuzzy clustering (CSW-WLIFC) as well as a proposed Whale Levenberg Marquardt Neural Network (WLM-NN) classifier is used in this paper to present a dietary evaluation system. The current WLI-FC algorithm is used in the CSW-WLIFC-based segmentation to split images. With segmentation accuracy, macro average accuracy, standard accuracy, and mean square error values of 0.999, 0.9643, 0.9627, and 0.0184, respectively, simulation results demonstrate that the suggested model performs better than current methods. J. Zhang et al. [22] experiments on the underwater datasets URPC2019 and URPC2020 show that their proposed framework not only alleviates the interference of underwater image degradation but also makes the mAP-0.5 reach 79.8% and 79.4% and improves the mAP-0.5 by 3.8% and 1.1%, respectively, when compared with the original YOLOv8 on URPC2019 and URPC2020, demonstrating that the proposed framework presents superior performance for the high-precision detection of marine organisms.

### 2.3 Comparison between existing works

In the previous papers, [9] mentioned in their experimental results, that food ingredients could be recognized with an accuracy of 82.02% using YOLOv3-tiny. To accomplish simultaneous multi-object recognition and adaptation, the author of [11] constructed a deep convolutional neural network coupled with YOLO, an advanced detection approach, and was able to achieve approximately 0.80% mean average precision(mAP). Furthermore, the detector of [12] was chosen to be the EfficientNet architecture, which provides fast and highly accurate detection as shown by a mAP of 0.92. X. Yin et al.[19] in their developed a detection framework combining PENet and YOLOv3 which is (PE-YOLO) and it achieved advanced results, achieving 78.0% in mAP and 53.6 in FPS.

J. Zhang et al. [22] experiments on the underwater datasets URPC2019 and URPC2020 show that their proposed framework not only alleviates the interference of underwater image degradation but also makes the mAP@0.5 reach 79.8% and 79.4% and improves the mAP@0.5 by 3.8% and 1.1%, respectively, when compared with the original YOLOv8 on URPC2019 and URPC2020, demonstrating that the proposed framework presents superior performance for the high-precision detection of marine organisms. On the other hand, [18] achieved in their Real-Time Flying Object Detection with YOLOv8 a mAP50 of 79.2%, mAP50-95 of 68.5%, and an average inference speed of 50 frames per second (fps) on 1080p videos.

Food-101, UECFOOD-256, and UECFOOD-100 are the three food datasets that [13] used a deep convolutional neural network (DCNN). Using 100 photos from each food category that were taken from each dataset, they investigated the efficacy of pre-training and fine-tuning a DCNN model. The evaluation revealed that the classification accuracy for UECFOOD-100, UECFOOD-256, and Food-101 was 78.77%, 67.57%, and 70.4%, respectively. Similarly, using the Food-101, UEC FOOD-100, and UECFOOD-256 datasets, the study by [14] deployed the Inception-v3 deep network created by Google [15]. The fine-tuned model V3 achieved a higher classification accuracy than the fine-tuned version of DCNN, i.e., 88.2%, 81.45%, and 76.17% for UECFOOD-100 and UECFOOD-256, and Food-101, respectively. Beijbom O. et al. [16] used a small-scale food dataset with around 100 training images for each food type to examine the effectiveness of deep convolutional neural network (DCNN) pretraining and fine-tuning. With the UEC-FOOD100/256 dataset, they obtained the highest classification accuracy in the experiments, 78.77% and 67.57%.

Table 2.3.1. Comparative analysis with existing work

| SL No | Author Name | Used Algorithm | Best Accuracy with Algorithm |
|---|---|---|---|
| 1. | Sahoo, et al. [24] | ResNet-50<br>ResNet-101<br>ResNeXt-50<br>SENet x ResNeXt-50 | SENet ResNeXt-50 = 83.2% |
| 2. | X. Yin et al. [19] | PE-YOLO (PENet + YOLOv3) | mAP: 78.0%, FPS: 53.6 |

| | | | |
|---|---|---|---|
| 3. | Dewantara et al. [9] | YOLOv3-tiny | Accuracy: 82.02% |
| 4. | Sun, J., et al. [11] | Deep CNN + YOLO | mAP: 0.80% |
| 5. | Liu et al. [12] | EfficientNet | mAP: 0.92 |
| 6. | J. Zhang et al. [22] | YOLOv8 (modified) | mAP@0.5: 79.8% (URPC2019), 79.4% (URPC2020) |
| 7 | D. Reis et al. [20] | YOLOv8 | mAP50: 79.2%, mAP50-95: 68.5%, FPS: 50 |
| 8. | Krizhevsky et al. [13] | DCNN | UECFOOD-100: 78.77%, UECFOOD-256: 67.57%, Food-101: 70.4% |
| 9 | Zhu F. et al. [23] | Inception-v3 | UECFOOD-100: 88.2%, UECFOOD-256: 81.45%, Food-101: 76.17% |
| 10 | Chen M. et al.[15] | DCNN | UECFOOD-100: 78.77%, UECFOOD-256: 67.57% |

### 2.4 Open Issues

Automatic food recognition has great assurance using computer vision but it faces several critical challenges. One significant issue is the limitation in recognition of complex images. Current models struggle with complex datasets that contain multiple objects/items from different backgrounds and angles.

The narrow range of available datasets is another difficulty. There are many models that rely heavily on training data or pre-trained models that do not represent the vast array of global items. This lack of viciousness can create problems for biased models that perform poorly on food items outside of the pre-trained model. Furthermore, there are many datasets that have isolated food items, the lack of variability of blurry images, messy backgrounds, and various lighting conditions. Also, there is not enough food quantity analysis in current research. For this reason, the lack of estimating the quantity of food consumed restricts the practical use of automatic food recognition systems.

We can solve these open issues but we would need to develop more robust algorithms that can detect food up to ingredient level. If we overcome these challenges, can fully utilize the potential of automatic food recognition in dietary assessment and personalized nutrition applications.

## 2.5 Summary

This chapter highlights the existing research on leveraging computer vision for automatic food recognition in images. This technology has the potential to revolutionize various fields, including dietary assessments. The main focus of this review was on deep learning techniques, especially YOLO, convolutional neural networks (CNNs), were the primary method for food identification. It compares the performance of different algorithms on various food image datasets. The highest accuracy noted is 88.2%, achieved with the Inception-v3 deep network. However, there's still a gap to bridge, especially in accurately recognizing complex dishes. Future research approaches could involve developing more powerful algorithms that can handle a wider variety of food images, containing various presentations and lighting conditions. Additionally, exploring techniques for accurate food quantity estimation within an image would further enhance the relevancy of this technology.

# CHAPTER 3

# RESEARCH METHODOLOGY/ REQUIREMENT ANALYSIS & DESIGN SPECIFICATION

## 3.1 Overview

This study utilizes a YOLO-based object detection model that is designed to automate the detection of food from random images. By integrating that system with personal information, we built a personalized recommendation system to help with the dietary management of diabetes during pregnancy.

Our concentration is on automated food detection, the model can identify food from images and classify specific food categories. The primary instrumentation for this study is YOLOv8, a state-of-the-art object detection algorithm. It is known for its efficiency and accuracy in real-time image processing. Our study concentrated on a distinct domain for the dataset, which is women with diabetes during pregnancy. By visiting some healthcare centers we gathered information on what kind of food they generally prescribe them. Based on their suggestion, we have collected food images from different contexts and angles with **18** food classes. These classes were chosen based on dietary guidelines and nutritional requirements. Our collected images were raw, images were not preprocessed or annotated. So it was up to us to finish the annotating process. After preparing the dataset by preprocessing, the dataset was split into three categories: training (80%), validation (10%), and test (10%) and prepared for training.

## 3.2 Proposed Methodology/ System Design

The proposed method is broken down into many stages. Figure 3.2.1 shows the whole process of our study.

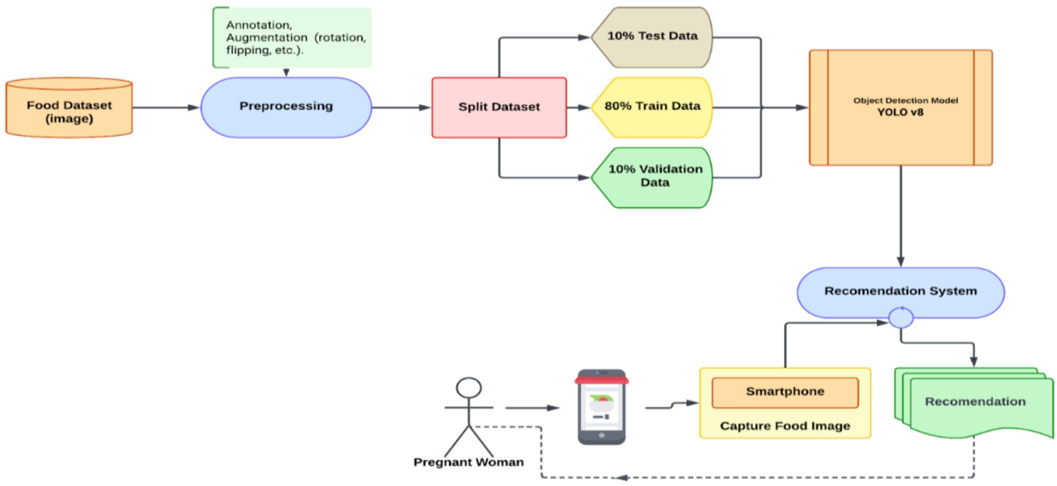



Figure 3.2.1: Proposed Methodology

First, a comprehensive dataset of food images has already been acquired. Following that, techniques for image processing such as annotation, augmentation, and data splitting are also applied to prepare the dataset for training. We have split the dataset into three categories: training (80%), validation (10%), and test (10%). Our goal is to detect food from images we used the YOLOv8, a state-of-the-art object detection algorithm known for its high detection accuracy. We also developed an Android-based application for the recommendation system.

### 3.2.1 Data Preprocessing

This section points out the process for the collection of data and emphasizes the value of having a broad and representative collection of food images. Our study concentrated on a distinct domain for the dataset, which is women with diabetes during pregnancy. We have collected a total of 5,016 images with **18** food classes of that domain. A diverse set of images with different contexts and angles were downloaded from various online sources like Google Images and social media food review pages etc. Along with the online source many images were captured manually through smartphones. These images were taken at home and from real-world scenarios.

#### 3.2.1.1 Dataset:

- **Total Images**: ~ 5,016 images.
- **Total Classes:** ~ 18 classes
- **Annotation Format**: YOLOv8-compatible (.txt files with bounding box coordinates).

Here are some samples of our dataset:


Figure 3.2.2: Sample Food Images of Dataset

### 3.2.1.2 Data Annotation

The data annotation process is one of the crucial processes in object detection. It is a mandatory process before training the YOLOv8 model. Accurate detection and classification of detected items also rely on the annotation step. Our collected images were raw, images were not preprocessed or annotated. So it was up to us to finish the annotating process.

To complete the annotation process we utilized the (**makesense.ai).** It is an open-source and free-to-use online tool for annotation images. We defined the label names and marked rectangular bounding boxes around the food items to manually annotate each of the dataset's 5,016 images.

Figure 3.2.3: Sample of Annotated Food Images with Bounding Boxes

After manually completing the annotation process of 5,016 images with corresponding labels, we exported the annotation file in **YOLO (.txt)** format which contains:

Table 3.2.1: Sample of Annotation format for YOLO (.txt)

| class_ID | center_X | center_Y | width | height |
|---|---|---|---|---|
| 1 | 0.360149 | 0.399367 | 0.146963 | 0.145570 |

### 3.2.1.3 Data Augmentation:

Data augmentation was required for our dataset because our initial dataset size was around 4000 images and there was a data imbalance in the number of images in each class. To make them balance and improve the model we applied some data augmentation techniques:

- Horizontal Flip (with a probability of 0.5)
- Random Brightness and Contrast (with a probability of 0.2)
- Rotation (random angle between -40 and 40 degrees with a probability of 0.5)

### 3.2.2 Resizing:

Data preprocessing is one of the crucial parts of any object detection model like YOLOv8. As we prepared a custom dataset all the images were raw. The dataset contains a total of 5,016 images. All the images were annotated using an online annotation tool (**makesense.ai).**

YOLO model patterns always require an input stream of fixed-size pictures. YOLOv8 architecture takes 640x640 pixel images as input before training the model. So, It was a mandatory part that we resize the images by 640x640 pixels. The image resizing process was not a manual process, instead, It was an automated process because YOLO offers a parameter "imgsz" that takes the image size like this **imgsz = 640** and automatically resizes the images before training the model.

| Original Image | YOLOv8 Resized - 640x640 |
|---|---|

Figure 3.2.4: Sample of the resized image using YOLOv8

### 3.2.3 Data Splitting

Figure 3.2.4 shows the preprocessed dataset was split into three categories: training (80%), validation (10%), and test (10%).

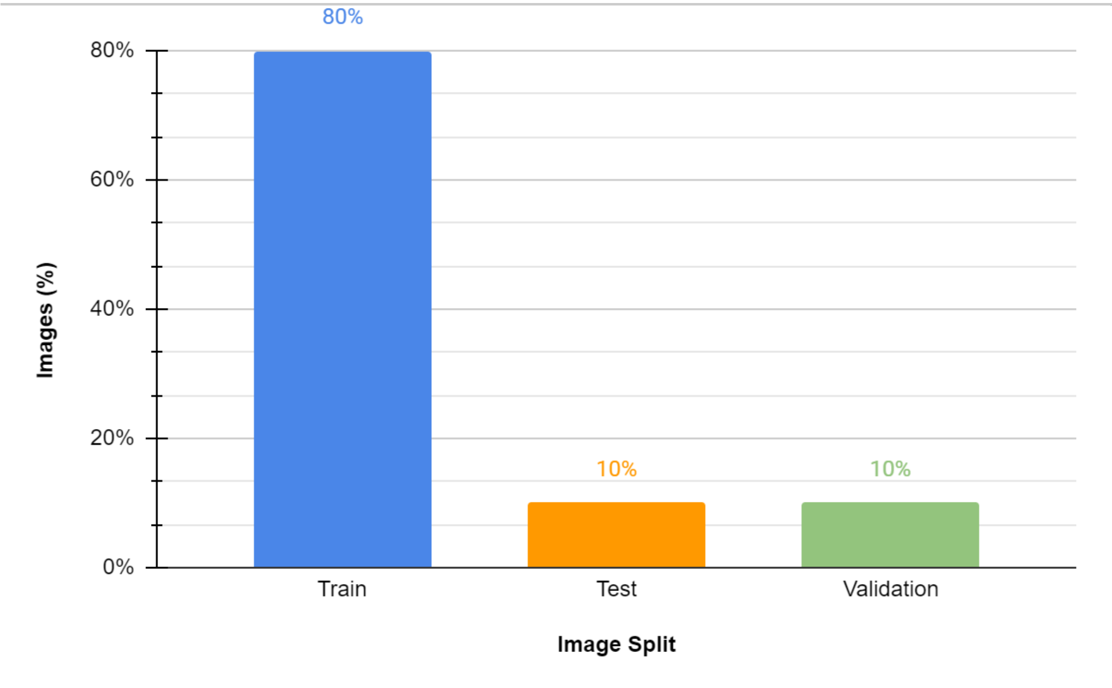



Figure 3.2.4: Dataset Split Train, Test and Validation

Number of Images in each partition:

- Training set (80%) ~ 4,006 images
- Test set (10%) ~ 513 images
- Validation set (10%) ~ 497 images

### 3.2.4 Model Selection

YOLOv8 is a state-of-the-art object detection algorithm known for its high detection accuracy. We chose these models (YOLOv8n and YOLOv8s variants as pre-trained weights). Because of its balance of high detection accuracy and computational efficiency. It is also suitable to implement on mobile applications so we chose this model.

#### 3.2.4.1 YOLOv8

YOLOv8 is a new state-of-the-art computer vision model built by Ultralytics. It works by generating a grid out of an image and using a single pass to predict bounding boxes and class probabilities for every grid cell. YOLOv8 can do both classification and detection together. With better detection precision and computational effectiveness over previous versions, YOLOv8 is a step ahead and is suitable for real-time applications. It can efficiently handle complex scenarios and a wide range of item sizes.

In the YOLOv8 family, there are multiple variants of models. The YOLOv8 models that were trained on the public dataset (coco) are shown here.

Table 3.2.2: Performance for YOLOv8 family models.

| **Model** | **Size** | **mAP50-95** | **Speed CPU ONNX (ms)** | **Speed A100 TensorRT (ms)** | **Params** | **FLOPs (B)** |
|---|---|---|---|---|---|---|
| YOLOv8n | 640 | 37.3 | 80.4 | 0.99 | 3.2 | 8.7 |
| YOLOv8s | 640 | 44.9 | 128.4 | 1.20 | 11.2 | 28.6 |
| YOLOv8m | 640 | 50.2 | 234.7 | 1.83 | 25.9 | 78.9 |
| YOLOv8l | 640 | 52.9 | 375.2 | 2.39 | 43.7 | 165.2 |
| YOLOv8x | 640 | 53.9 | 479.1 | 3.53 | 68.2 | 257.8 |

Yolov8 comes with several features such as:

- It comes with a new backbone network
- Now it is easy to compare model performance with older models in the YOLO family
- Introduced a new loss function and
- A new anchor-free detection head

### 3.2.4.2 YOLOv8 Architecture

YOLOv8 is a state-of-the-art object detection mode. The architecture design of this model is modular and scalable. The model architecture has three main components: (1) Backbone, (2) Neck, and (3) Head.

The Backbone is used to extract features using a convolutional neural network (CNN) from the input image. YOLOv8 uses a custom backbone called CSPDarknet53. The Neck, also known as a feature extractor, is used to manage feature maps of different stages of the backbone to gather information at various scales. YOLOv8 utilizes a C2F module instead of FPN. The head is for the predictions. It uses multiple detection modules to predict the bounding boxes, objectless score, and class probabilities. Figure 3.2.5 shows the YOLOv8 architecture:

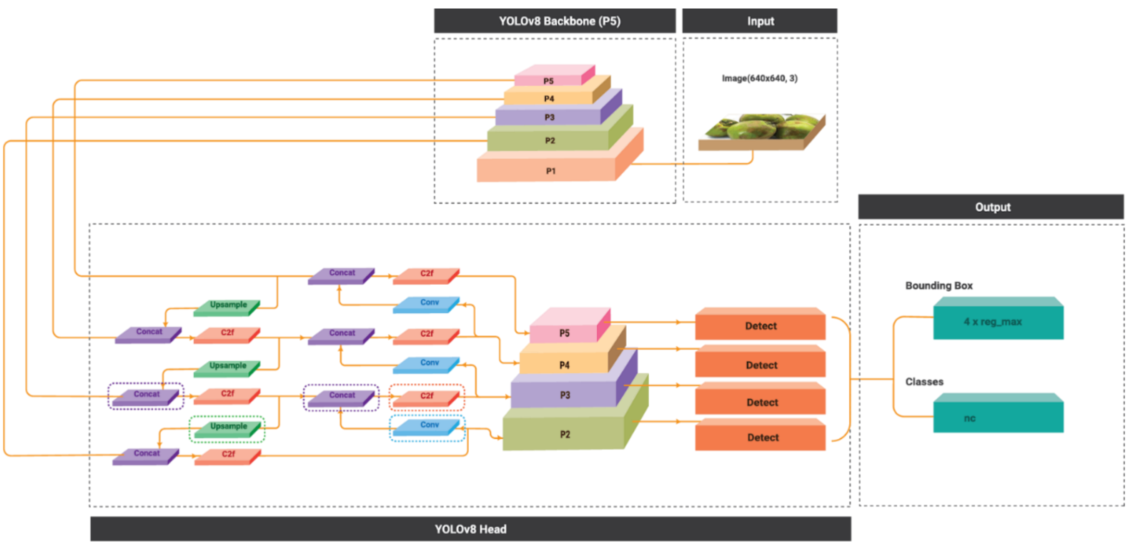



Figure 3.2.5: YOLOv8 Architecture

In our project, YOLOv8n (the nano model) and YOLOv8s (the small model) were used as the pre-trained weights with our custom dataset. YOLOv8 has the nano to extra large model, where the nano and small models are good for small datasets and for low-end devices to detect objects accurately. As we have a comparatively small dataset we used the YOLOv8n and YOLOv8s models, based on the performance we use the YOLOv8s as the final pre-trained weight, which also helps us when implementing it for mobile applications to develop the recommendation system.

## 3.3 Hardware/ Software Requirement

In the development and training of the model using the pre-trained model YOLOv8n and YOLOv8s and the implementation of the recommendation system, a combination of software, core libraries (Ultralytics YOLO, TensorFlow, Keras, OpenCV, Matplotlib, NumPy, Pandas, Scikit-Learn, etc.), local hardware and cloud-based resources was utilized:

**3.3.1 Hardware Specifications (Cloud) -** Training purpose

**Environment:** Kaggle Notebooks, Google CoLab

**GPU**: NVIDIA Tesla P100 GPU (16GB Dedicated Memory)

**RAM**: Up to 30 GB

**Disk:** 73 GB Storage

**3.3.2 Hardware Specifications (Local & Cloud) -** Preprocessing purpose

**Environment:** VS Code

**CPU**: AMD Ryzen 5 2600X, 3.6 GHz

**RAM**: 16 GB

**GPU:** AMD Radeon RX 580

**Storage**: 1 TB SSD

**3.3.3 Software Specifications (**Recommendation System**) -** Development purpose

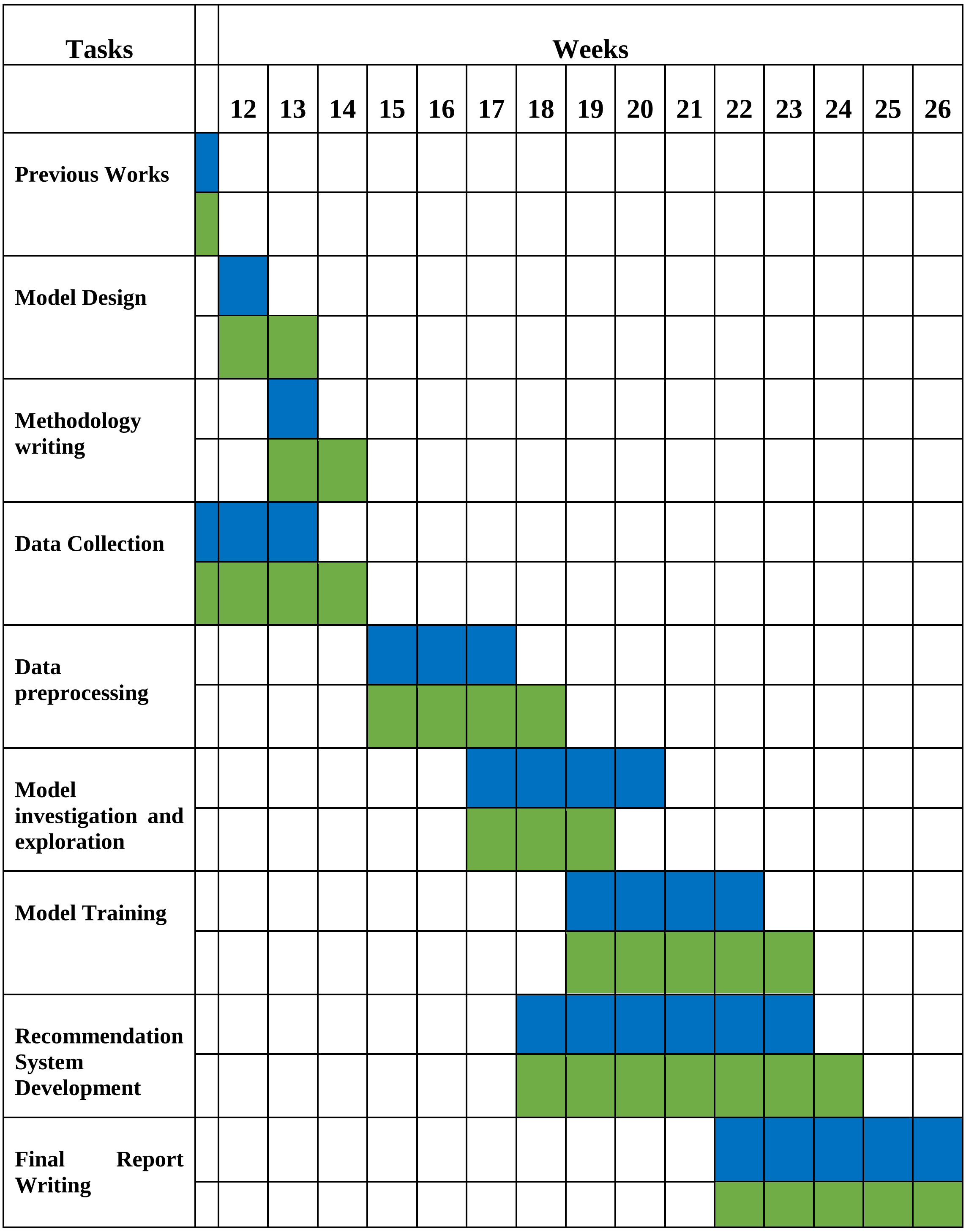
**Environment:** Android Studio Jellyfish 2023
**Database:** Firebase

## 3.4 Project Management and Financial Analysis

### 3.4.1 Project Management

This Gantt chart highlights progress made and indicates adjustments over the last five months.

Table 3.4.1.Gantt Chart of Project Management



| | |
|---|---|
| **Estimated Work Period** | |
| **Actual Work Period** | |

The table outlines the last five months' research project timeline that progresses from theoretical preparation to the final report. The process begins with creating a working plan, which is a continuous task that is done throughout the project's lifecycle. Previous works show the time of the working plan. From model investigation to Final report writing is the main working period before writing the report.

### 3.4.2 Financial Analysis

Table 3.4.2 Financial Analysis

| SN | Components | Estimated Cost (BDT) |
|---|---|---|
| 01 | Research (data scientists, machine learning engineers) | 20,000-50,000 |
| 02 | Servers, cloud computing resources, data storage systems | 45,000-70,000 |
| 03 | Travel expenses for meetings, presentations, etc. | 3000-5000 |
| 04 | Acquiring and annotating food image dataset | 10,000-20,000 |
| 05 | High-performance GPUs, development tools, software licenses | 160,000-280,000 |
| 06 | User testing, clinical trials, collaboration with healthcare professionals | 18,000-25,000 |
| 07 | App development, software updates | 80,000-120,000 |
| 08 | Technical documentation, project reports, etc. | 10,000-15,000 |
| 09 | Additional tools and miscellaneous costs | 10,000-15,000 |
| 10 | Buffer for unforeseen expenses and adjustments | 5000-10,000 |
| | **Total Estimated Cost** | **3,61,000BDT -6,10,000 BDT** |

### 3.5 Summary

This study's main concentration is automated food detection, the model can identify food from images and classify specific food categories. The methodology consists of gathering a variety of image datasets, using image processing techniques. The creation of a recommendation system, and partitioning the dataset are emphasized. Real-time, online data collection is done with an emphasis on health and nutrition statistics. Project management ensures a strong application of cutting-edge technology and risk reduction by emphasizing objectives, timeliness, and financial concerns.

# CHAPTER 4
# IMPLEMENTATION

## 4.1 Overview

To train our model we have utilized YOLOv8's nano and a small pre-trained model with our custom dataset. It is a lightweight object detection model ideal for our dataset. It is known for its high accuracy and efficiency. We have trained our model in a cloud platform that provides free GPU and sufficient memory to train the object detection model. The training procedure was started by structuring the files in YOLOv8 format and loading the custom preprocessed dataset. Both nano and small pre-trained weights were utilized to train the model on our dataset to inspect which one brings better results. We have trained the models many times with different hyper parameters (such as epochs, batch, learning rate, early stopping, etc.) To get good accuracy and optimize the training process.

During the training cycle, a validation evaluation is also performed with a validation set. By evaluating that it was easy to observe the model performance by reviewing the recall-precision values and mean average precision (mAP). To evaluate the model performance a test set is also employed after training. To evaluate the model performance with real-life entities an inference was run over random images that the model had not seen before. This type of method helped to test the model before implementing it in the recommendation system. Finally, an app was also developed to implement our model to detect objects by capturing food images.

By combining the accurate food detection capabilities of the YOLOv8s model with the mobile application, the system achieves the expected functionality.

## 4.2 Train Model/ Prototype Design

### 4.2.1 File Structure

To train our model first of all we have prepared the dataset and structured it in YOLOv8 format. The file structure is like this:

```
|------- Project Folder
            |           |
            |           |-------images (image files class wise)
            |           |           |
            |           |           |----- test
            |           |           |----- train
            |           |           |----- val
            |           |
            |           |
            |           |-------labels (annotation .txt files class wise)
            |                       |
            |                       |----- test
            |                       |----- train
            |                       |----- val
            |
            |-------- dataset.yaml
```

Figure 4.2.1: File Structure before training in YOLOv8 format

### 4.2.3 Hyper parameter Tuning

Hyper parameter tuning is another important part of model training. In our projects, we have used some hyper parameters to optimize the model performance.

- **Batch size:** Batch size is the number of samples processed during the training. In YOLOv8 the default batch size is 64 which requires more memory. But this size is changeable. As we used cloud hardware with GPU acceleration and 16GB memory to train the model, we tested the model with different batch sizes. Initially, we used batch size 8 which made the training process very slow. After finishing the process run the model again with a higher batch size of 16, 32, and 64 respectively. GPU acceleration reduced the time of the training process and we completed several epochs (20, 30, 50, 100) and got the evaluation metrics.
- **Epochs:** Initially, the YOLOv8 model was trained up to 100 epochs with different batch and image sizes. But our final epoch size was 50 with batch size 64 and image size 640. The final epochs were set after testing the model with different epoch sizes, batch sizes, and image sizes (640) after several time training processes and evaluating the result.
- **Patience:** This hyper parameter helps the model prevent overfitting during the training process. When the model runs the validation set keeps improving, if the model performance does not improve a specific number of epochs/iterations it stops the process. This is called early stopping. It assists to prevent overfitting.

### 4.2.4 Food Detection Training Process

The process starts by splitting our custom dataset into test, train, and validation. After that put those in a file structure that YOLOv8 supports. The next step was loading pre-trained **YOLOv8n and YOLOv8s (**one after finishing another) weights using Ultralytics to start training the model. YOLOv8 nano and small were pre-trained on massive datasets containing various objects which is a strong foundation for object detection tasks with custom datasets. Along with that as our data were prepared we also trained the **YOLOv5** model using a pre-trained weight (YOLOv5n).

As we used a custom dataset and two models two **.yaml** files were required of their format that we created in a separate file providing the train, test, and validation file path and **18 class** names to classify the food names after detection. These models, like YOLOv5n, YOLOv5s, YOLOv8n, and YOLOv8s, are pre-trained on massive datasets containing various objects.

In the training loop, the dataset goes through an intensive process. First, the input images are preprocessed like resizing to 640x640, as the YOLOv8 and v5 accept this size as input and normalize the pixel values. Next step, the preprocessed images in fed to the YOLO model architecture, which generates a prediction for bounding boxes and class probabilities.

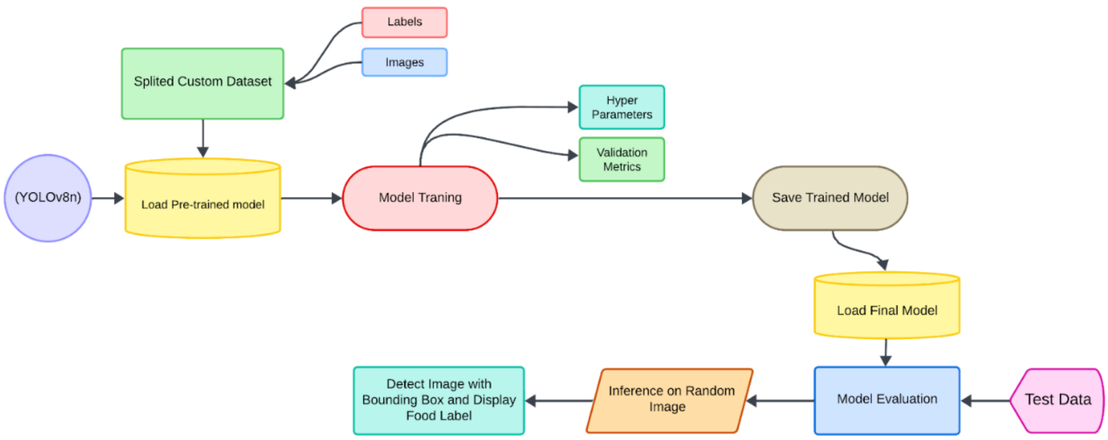



Figure 4.2.2: YOLOv8 food detection workflow

After that, the model goes through several other processes such as loss calculation, weight adjusting to the model's internal layer to minimize the loss in the next iteration, etc. Also during the model training, some hyper-parameters are used to optimize the model performance. Furthermore, a separate validation set is also used to periodically evaluate the model performance during training, it also helps to prevent overfitting.

After training the model their best weight **(best. pt)** was saved in a separate folder. To examine the trained model and how it performs, again we have loaded the model (this time the best weight) and used a test set to evaluate the model performance. And finally, we run the inference on a random image that the model hasn't seen before. After visualizing the detected image we found that the model used its learned knowledge to detect the food item from the image with proper bounding boxes and classes.

## 4.3 System Testing/ Model Evaluation

### 4.3.1 Evaluation Metrics

To evaluate the model detection capabilities to our model, several performance metrics were used to calculate the values,

- **Precision:** It measures the accuracy of positive predictions.

$$1. Precision = \frac{TP}{\mathrm{TP} + \mathrm{FP}}$$

- **Recall:** It defines the ability of a model to find positive instances and avoid false negatives.

$$2. Recall = \frac{TP}{\mathrm{TP} + \mathrm{FN}}$$

- **Intersection over Union (IoU)**: It shows how well the ground truth box and the predicted bounding box overlap.

$$3. IoU = \frac{Area\ of\ Overlap}{\text{Area of Union}}$$

- **Mean Average Precision (mAP50)**: Mean average precision calculated at an intersection over a union (IoU) threshold of 0.50. It assesses the model performance.

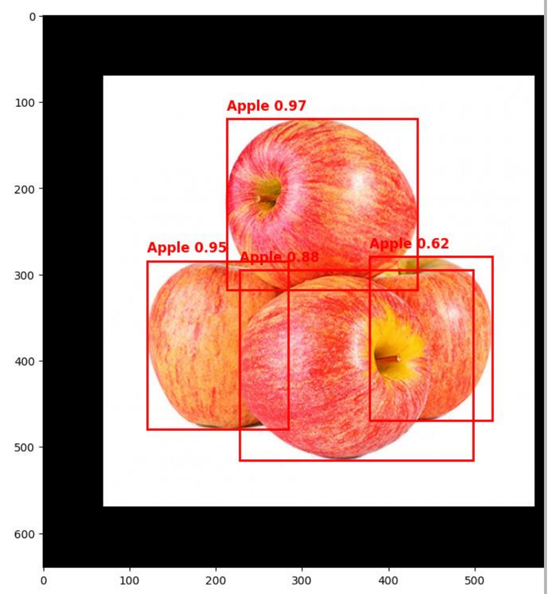
- **F1 Score:** The F1 score is the harmonic mean of precision and recall. A higher F1 score indicates better overall performance.

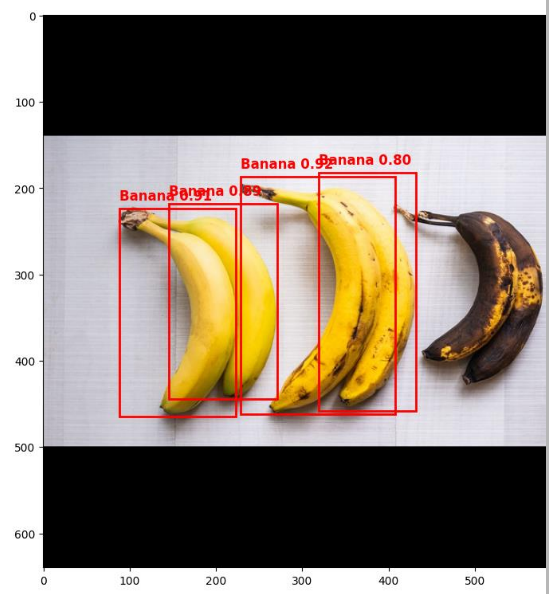
$$4. F1\ Score = \frac{2 * (Precision * Recall)}{\text{Precision} + \text{Recall}}$$

- **Mean Average Precision (mAP50-95)**: mAP50-95 considers a wider range of IoU thresholds into account (range from 0.50 to 0.95), and it offers a more complete assessment that captures both high and low overlap between predicted and ground truth bounding boxes.

### 4.3.2 Model Generalization Testing

To evaluate the model performance, we also performed a test evaluation using a test set to evaluate the model performance. In the test evaluation process precision, recall, mAP50, and mAP50-95 all metrics were used to get the score of the model performance. Finally, we ran the inference on a random image that the model hadn't seen before, several random images of each class were used to perform the inference. All the inferences on images ran very well and got accurate detection of food items. Figure 0.0.0 shows the inference testing,




| | |
|---|---|
| **Size:** 640x640<br>**Detection:** 4 Apples, 712.1ms<br>**Speed:** 4.0ms preprocess, 712.1ms inference, 1.8ms post-process per image at shape (1, 3, 640, 640) | **Size:** 640x640<br>**Detection:** 4 Bananas, 650.5ms<br>**Speed:** 2.4ms preprocess, 650.5ms inference, 1.0ms post-process per image at shape (1, 3, 640, 640) |

Figure 4.3.1: Sample of inference results on a random image

### 4.3.3 Mobile App Functionality Testing

As planned our best model was also implemented on mobile applications which is **(mDiet)**. After completing the training process with two pre-trained weights we got our best-trained model with evaluation metrics. Observing the evaluation metrics we found our model trained with YOLOv8s (small) got better accuracy **(mAP50 - 0.882)** than YOLOv8s and YOLOv5s. To use this model on Android we had to convert the pyTorch file (best. pt) to a TensorFlow Lite file which is (best. tflite). So using the Ultralytics we converted the file to tflite properly. Files are **(best_float16.tflite and best_float32.tflite).** This tflite model was used in our mobile application to run inference on captured or uploaded images. After capturing and uploading some raw images we found that the model was working properly by detecting accurately and visualizing the bounding boxes with a confidence score. Some testing samples are shown here in Figure 4.3.2,

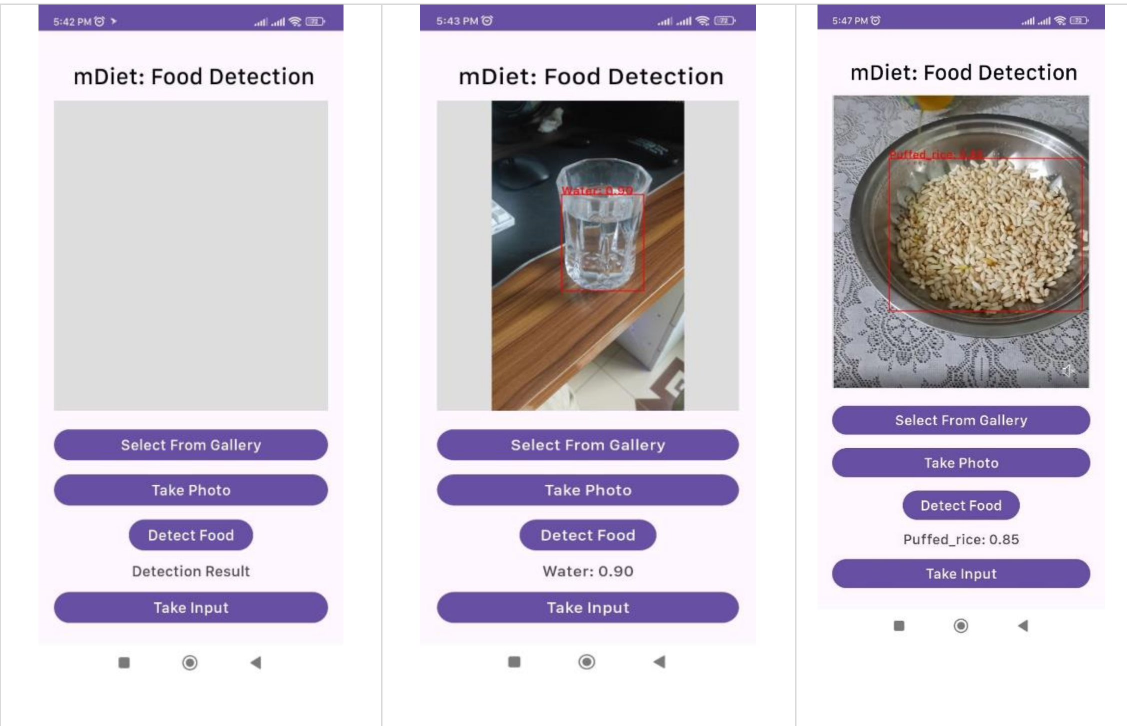



Figure 4.3.2: Sample of detection on mobile application

### 4.4 Summary

This portion summarizes the training of a YOLOv8 model. YOLOv5n, v5s, YOLOv8n, and YOLOv8s models are efficient for comparatively small datasets like ours. These models are evaluated and trained in a cloud environment with GPU acceleration. Hyper parameters were carefully tuned to balance training speed and accuracy. The model's performance was precisely evaluated using unseen data and metrics like mAP (mean Average Precision) to ensure the model performance. After observing successful training, the best-performing YOLOv8s model was converted for mobile deployment. This allows the mobile app to capture food images and utilize the model to detect the food items within them, feeding valuable information to the recommendation engine.

# CHAPTER 5
# RESULT AND ANALYSIS

## 5.1 Overview

This section presents the experimental results from training and evaluating food detection models, as discussed in Chapter 4, including their integration into a mobile application. Using YOLOv8 as the foundation, we trained with a custom dataset of 5,016 images across 18 classes, alongside YOLOv5 for comparison. Both models were trained over 50 epochs with a batch size of 64. YOLOv5 achieved a precision of 0.872, recall of 0.78, mAP50 of 0.851, and mAP50-95 of 0.651, indicating good accuracy. In contrast, YOLOv8 outperformed with a precision of 0.892, recall of 0.814, mAP50 of 0.882, and mAP50-95 of 0.709, demonstrating superior performance. Figures illustrate the reduction in box loss, class loss, and Distribution Focal Loss (DFL) for YOLOv8, confirming effective training and improved bounding box prediction and classification. The precision-recall curve further validates YOLOv8's accuracy in detecting food items, establishing it as the more accurate model for this task.

### 5.2.1 Experimental/ Simulation Result

This section shows the experimental results obtained from training and evaluating the model of food detection as discussed in Chapter 4 including the mobile application. As discussed, we use YOLOv8 as our foundation to train the model, we have used a custom dataset consisting of 5,016 images with 18 classes. With the preprocessing technique, we prepared the dataset for training. As discussed we have the YOLOv8 as the foundation but in the training time, we also have trained the YOLOv5 mode. So, a total of 2 models were trained on the training process, this section also highlights the experimental result of that mode.

Table 5.1.1 summarizes the performance result of the YOLOv5 model that was run up to 50 epochs with 64 batches.

Table 5.2.1 Model performance result of the YOLOv5n

| Class | Images | Instances | Box (P) | R | mAP50 | mAP50-95 |
|---|---|---|---|---|---|---|
| all | 497 | 2016 | 0.872 | 0.78 | 0.851 | 0.651 |
| Apple | 30 | 104 | 0.956 | 0.894 | 0.971 | 0.857 |
| Banana | 30 | 181 | 0.877 | 0.748 | 0.878 | 0.573 |
| Beef | 30 | 284 | 0.746 | 0.532 | 0.678 | 0.34 |
| Bread | 30 | 64 | 0.689 | 0.553 | 0.629 | 0.397 |
| Chicken | 30 | 198 | 0.774 | 0.707 | 0.789 | 0.535 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Coconut | 25 | 115 | 0.773 | 0.817 | 0.861 | 0.646 |
| Egg | 30 | 160 | 0.953 | 0.938 | 0.982 | 0.741 |
| Fish_Curry | 19 | 48 | 0.632 | 0.521 | 0.682 | 0.485 |
| Lemon | 26 | 126 | 0.952 | 0.857 | 0.953 | 0.756 |
| Mango | 30 | 156 | 0.927 | 0.885 | 0.944 | 0.828 |
| Milk | 29 | 47 | 0.892 | 0.877 | 0.903 | 0.769 |
| Noodles | 30 | 37 | 0.94 | 0.784 | 0.822 | 0.683 |
| Orange | 30 | 157 | 0.882 | 0.907 | 0.952 | 0.725 |
| Papaya | 18 | 172 | 0.921 | 0.849 | 0.904 | 0.639 |
| Puffed_rice | 20 | 23 | 1 | 0.944 | 0.985 | 0.89 |
| Rice | 30 | 52 | 0.914 | 0.538 | 0.614 | 0.471 |
| Soup | 30 | 33 | 0.98 | 0.879 | 0.924 | 0.78 |
| Water | 30 | 59 | 0.887 | 0.802 | 0.856 | 0.605 |

Initially, after running the YOLOv5 model it generated an evaluation result. Where it shows the overall performance of the model and the performance of each class individually. From the table, we can see that YOLOv5 with a pre-trained weight and custom dataset has precision **(P = 0.872, R = 0.78, mAP50 = 0.851 and mAP50-95 = 0.651)** which is quite good accuracy.

Table 5.1.2 summarizes the performance result of the YOLOv8 model that was run up to 50 epochs with 64 batches.

Table 5.2.2 Model performance result of the YOLOv8s

| **Class** | **Images** | **Instances** | **Box (P)** | **R** | **mAP50** | **mAP50-95** |
|---|---|---|---|---|---|---|
| all | 497 | 2016 | 0.892 | 0.814 | 0.882 | 0.709 |
| Apple | 30 | 104 | 0.916 | 0.938 | 0.986 | 0.905 |
| Banana | 30 | 181 | 0.865 | 0.781 | 0.88 | 0.565 |
| Beef | 30 | 284 | 0.801 | 0.676 | 0.776 | 0.42 |
| Bread | 30 | 64 | 0.767 | 0.72 | 0.765 | 0.505 |
| Chicken | 30 | 198 | 0.838 | 0.798 | 0.83 | 0.602 |
| Coconut | 25 | 115 | 0.917 | 0.783 | 0.87 | 0.719 |
| Egg | 30 | 160 | 0.98 | 0.906 | 0.981 | 0.779 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Fish_Curry | 19 | 48 | 0.791 | 0.711 | 0.775 | 0.599 |
| Lemon | 26 | 126 | 0.949 | 0.841 | 0.955 | 0.83 |
| Mango | 30 | 156 | 0.954 | 0.931 | 0.972 | 0.853 |
| Milk | 29 | 47 | 0.923 | 0.894 | 0.938 | 0.868 |
| Noodles | 30 | 37 | 0.944 | 0.784 | 0.836 | 0.726 |
| Orange | 30 | 157 | 0.903 | 0.888 | 0.962 | 0.778 |
| Papaya | 18 | 172 | 0.933 | 0.889 | 0.948 | 0.731 |
| Puffed_rice | 20 | 23 | 1 | 0.891 | 0.974 | 0.892 |
| Rice | 30 | 52 | 0.797 | 0.558 | 0.647 | 0.509 |
| Soup | 30 | 33 | 0.933 | 0.845 | 0.923 | 0.813 |
| Water | 30 | 59 | 0.843 | 0.814 | 0.864 | 0.667 |

From the table, we can see that YOLOv8 with a pre-trained weight and custom dataset has precision **(P = 0.892, R = 0.814, mAP50 = 0.882 and mAP50-95 = 7.09)** which is our best accuracy.

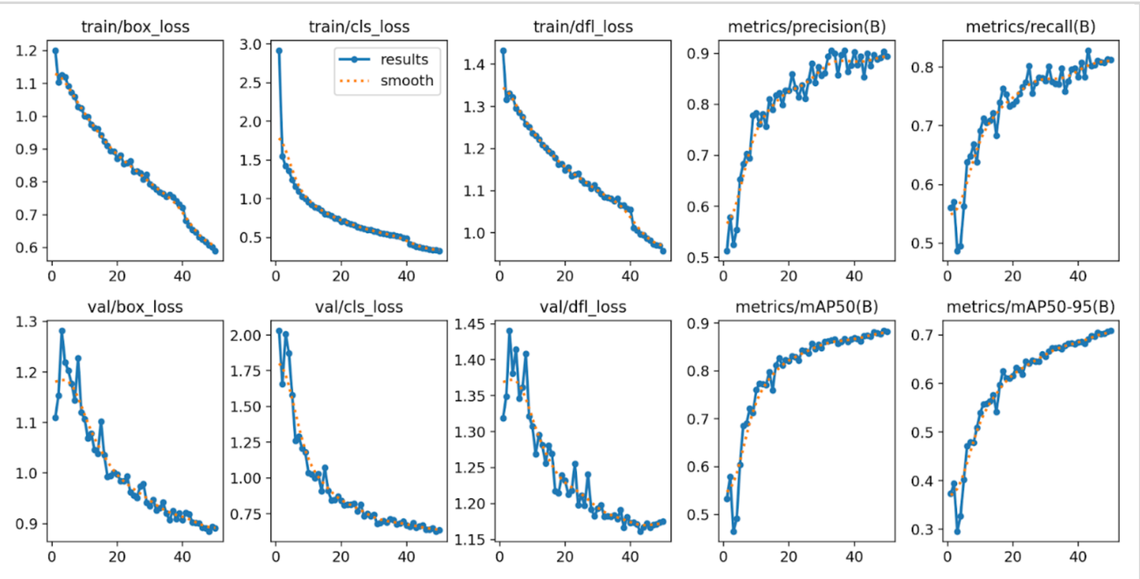



Figure 5.2.2: YOLOv8s losses and metrics

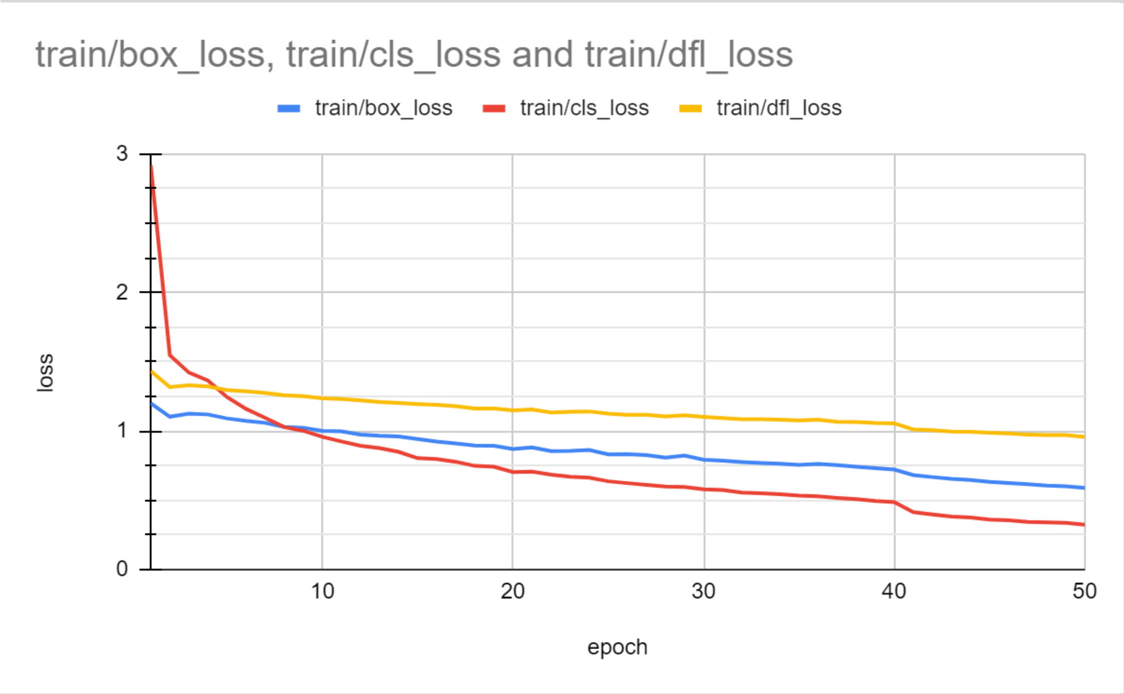



Figure 5.2.3: YOLOv8s Train, Training box_loss, class_loss, and dfl_loss

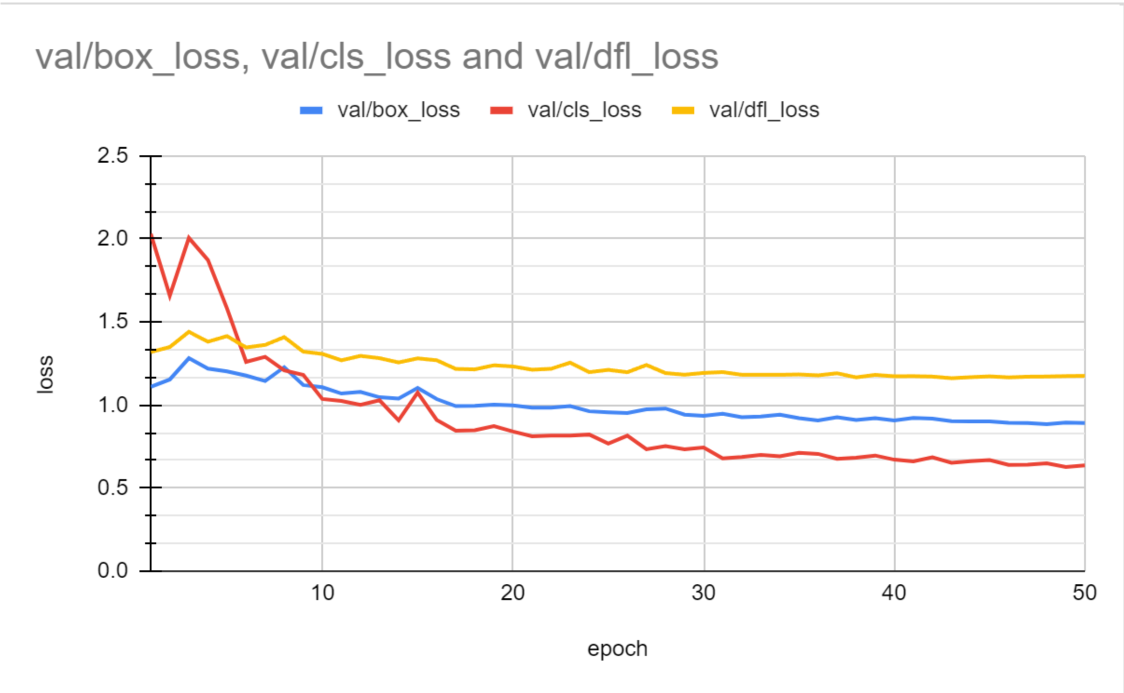



Figure 5.2.4: YOLOv8s Train, Validation box_loss, class_loss, dfl_loss

Figures 5.2.3 and 5.2.4 graph show the training and validation progress of the YOLOv8s model with a custom dataset over 50 epochs, following three loss components (train/val): box_loss, class_loss, and DFL (Distribution Focal Loss).

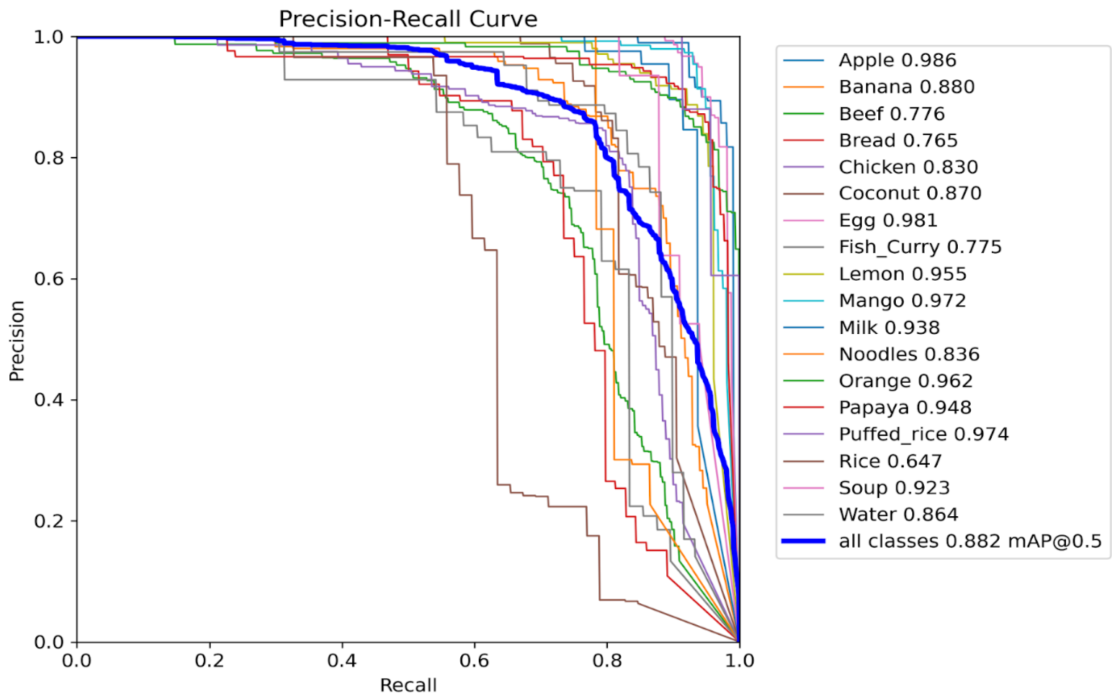



Figure 5.2.5: YOLOv8s Training Precision-Recall Curve

**Box loss:** As we can see this loss steadily decreases in both the train and validation set, meaning that the model improves in accurately predicting bounding box locations.

- **Class loss:** From the graph, we can see class loss was initially more elevated but it dropped rapidly in the early epochs which showed significant improvements in the model's ability to classify objects correctly.
- **dfl loss:** This decreases slowly, it reflects the improvements in the distribution focal aspect and contributes to better localization of predicted boxes.

Overall, all loss types suggest effective training, with the model becoming more accurate and exact in detection over time.

**Figure 5.2.2** shows the final YOLOv8s models' final precision-recall curve, evaluated on the validation set.

Overall we have used two models YOLOv5 and YOLOv8 for our study, YOLOv8 got the highest detection accuracy with a mean Average Precision (mAP50 = 0.882) and (mAP50-95 = .709) score which are the maximum among them. This value indicates that our model can detect food items from images.

### 5.2.2 Experiment on Mobile Application (mDiet)

After training the model and evaluating the performance score we have got our best mode. We plan to implement it on mobile applications. To make a mobile application an Android application building platform Android Studio was used. We have included three features in our application "Selecting Image From Gallery", "Image Capture", and "Recommendation System". To implement our model which is in the PyTorch version, we had to convert it to the TensorFlow Lite version (tflite). In the file mode, our Input Shape was [1, 640,640, 3], where 1 represents 1 image inference at a time as it is a small device, 640 is our image size and 3 indicates the RGB image. The output shape [1, 22, 8400] where indicates 22 predictions for each of the 8400 grid cells in the image. To decode this output shape there is no official code/algorithm, so we used the Non-Maximum Suppression (NMS) algorithm and thresholding to decode the output shape and extract valuable information such as food labels, bounding box coordinates, conf. Score etc. Here are some experimental result samples:

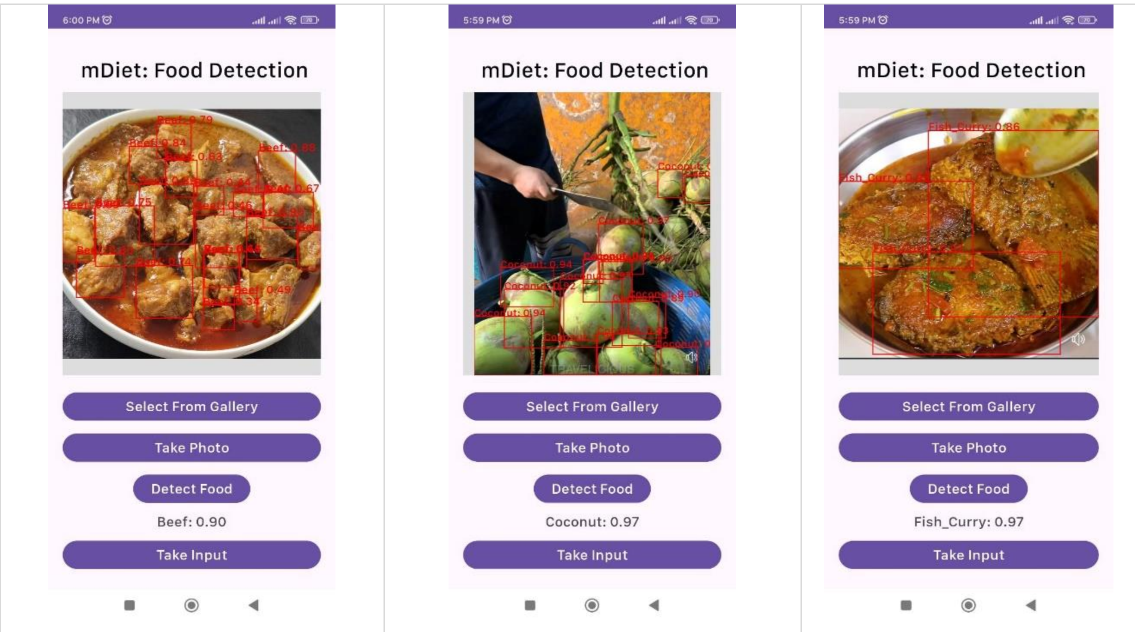



Figure 5.2.6: Experimental detection result on mobile application

After that, we have our recommendation system. This system is connected to the Firebase database. For real-time applications, firebase is one of the best options. We have put several pre-defined suggestions based on our food classes and some parameters related to pregnant women with diabetes. These suggestions will be provided on the user screen when they use the recommendation system. To do that they have to insert some input related to her pregnancy and health. Based on this input and the detected food class our system will generate a suggestion for the user. Here is a sample of our recommendation system:

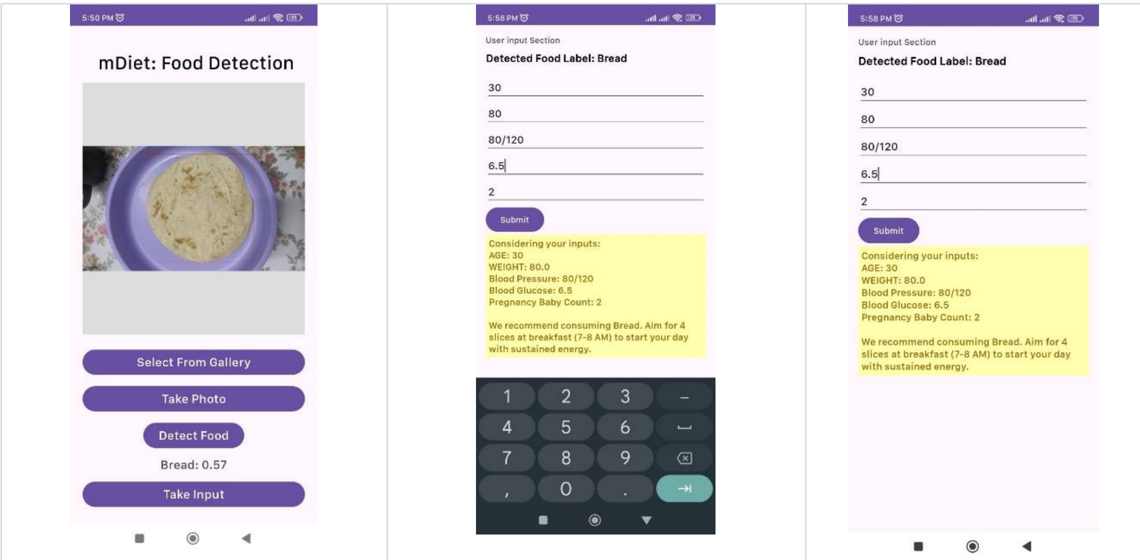



Figure 5.2.7: Sample of detection and Recommendation

## 5.3 Performance/ Comparative Analysis

### 5.3.1 Performance Comparison

Table 5.3.1 Model performance result of the trained models

| Model | P | R | mAP50 | mAP50-95 |
|---|---|---|---|---|
| YOLOv5s | 0.886 | 0.818 | 0.877 | 0.690 |
| YOLOv5n | 0.872 | 0.780 | 0.851 | 0.651 |
| YOLOv8n | 0.846 | 0.796 | 0.857 | 0.655 |
| YOLOv8s | 0.892 | 0.814 | 0.882 | 0.709 |

Table 5.3.2 Model performance result of the trained models

| Model | Training Box Loss | Training Class Loss | Training DFL Loss | Validation Box Loss | Validation Class Loss | Validation DFL Loss |
|---|---|---|---|---|---|---|
| YOLOv5s | 0.67212 | 0.41153 | 0.99698 | 0.92286 | 0.6307 | 1.1574 |
| YOLOv5n | 0.815 | 0.64057 | 1.08519 | 0.96689 | 0.78692 | 1.17264 |
| YOLOv8n | 0.80952 | 0.67756 | 1.10173 | 0.92969 | 0.776 | 1.1691 |
| YOLOv8s | 0.58943 | 0.32415 | 0.95686 | 0.89108 | 0.63668 | 1.17543 |


Figure 5.3.1: Trained model comparison by losses

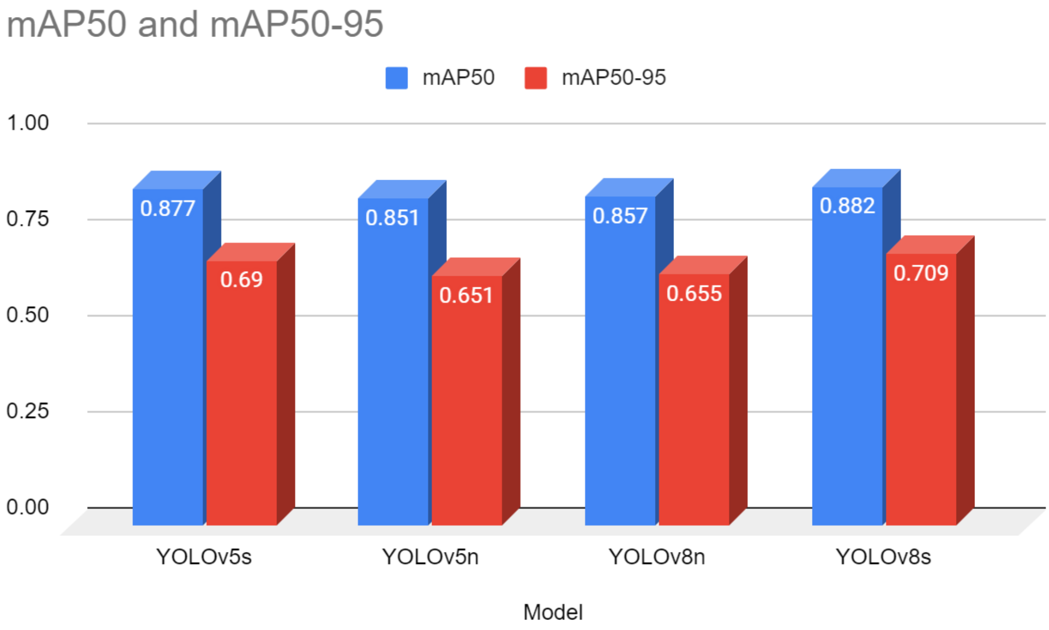



Figure 5.3.2: Trained model comparison by mAP50 and mAP50-95

In assessing models for food detection, **YOLOv8s appears as the top performer**, boasting the highest precision (89.2%) and recall (81.4%), along with superior mAP50 (88.2%) and mAP50-95 (0.709) scores. This shows its exceptional capability in accurately detecting food items, even in various conditions. YOLOv8n follows closely, with little lower metrics but still showing strong performance. In contrast, YOLOv5, while reliable with a precision of (87.2%) and recall of (78%), shows lower mAP scores, particularly at varied IoU thresholds, highlighting its relative limitations in complex detection scenarios. Overall, YOLOv8s is the most suitable choice for applications prioritizing high accuracy in food detection.

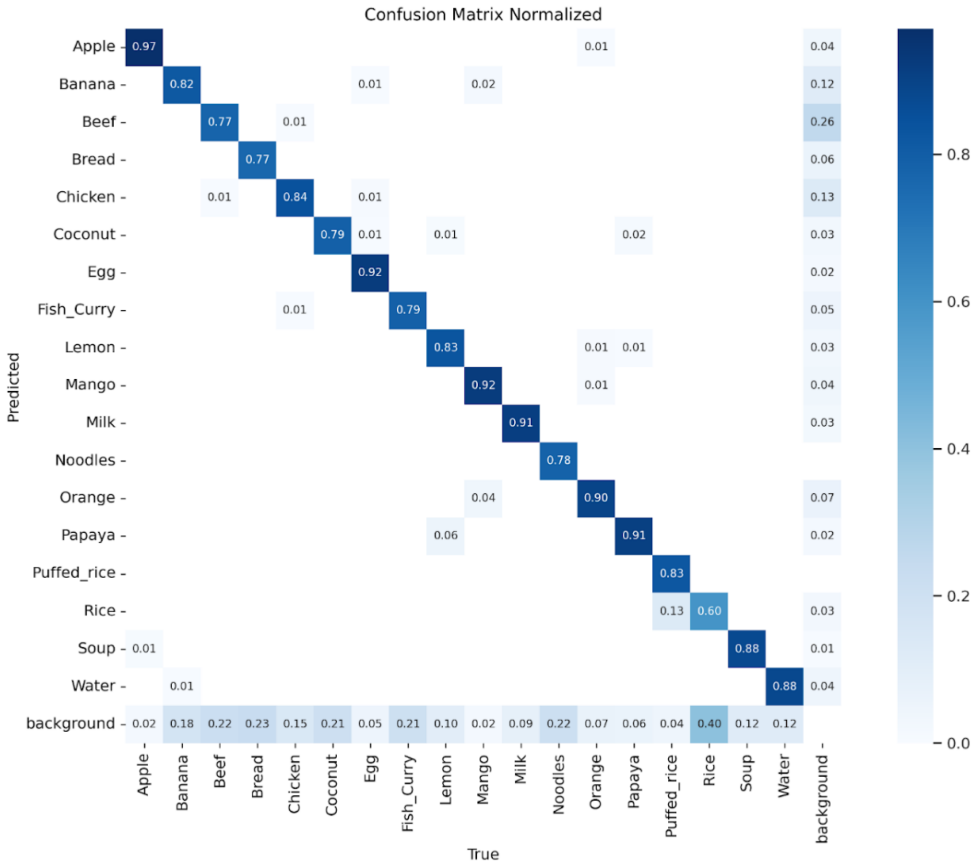



Figure 5.3.2: YOLOv8 Confusion Matrix (Normalized)

### 5.3.2 Application Performance

To implement the model in mobile we have converted it to **.tflite** format. Which is a lightweight version of our main model but with a similar or slightly different performance. To make the inference on the image faster we have used the (float16) version of tflite. I can complete the inference within mile seconds and detect the food items accurately.

## 5.4 Summary

In this chapter, the experimental results are discussed from training YOLOv8 and YOLOv5 models for food detection in detail. Using a custom dataset of 5,016 images with 18 classes, YOLOv8 showed outstanding performance with higher precision, recall, and mAP scores compared to YOLOv5. Both models were trained for 50 epochs with batch sizes of 64. YOLOv8's enhanced accuracy made it the ideal choice for food detection applications.

The chapter also discusses the integration of the YOLOv8 model into a mobile application (mDiet), converted to TensorFlow Lite for efficient on-device inference. The app features image capture and a recommendation system, connected to Firebase, offering dietary suggestions based on user inputs and detected food items. This system is particularly tailored for pregnant women with diabetes, providing personalized dietary advice. Overall, the integration of YOLOv8 in the mobile app showcases its effectiveness in practical, real-world scenarios, significantly enhancing user experience with accurate food detection and relevant recommendations.

# CHAPTER 6

# IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

## 6.1 Impact on Life

Technology may significantly improve the health of pregnant mothers and their unborn children by offering personalized and targeted dietary food recommendations. Through individualized dietary recommendations, technology can assist pregnant women in maintaining healthy blood glucose levels, hence reducing the occurrence of gestational diabetes mellitus and its associated complications.

When a pregnant mother faces Diabetes, she needs to be on a proper diet and the method helps to improve the precision and efficiency with which healthcare professionals can track and oversee pregnant women's food.

By using computer vision models like YOLOv8, the technology may ensure accurate food item detections, and suggest immediate individualized dietary recommendations are beneficial for pregnant women as they encourage appropriate food consumption. Pregnant women can easily record and evaluate their meals with the system's user-friendly interface, which provides them with the knowledge and tools to make dietary food choices. Additionally, by automating the dietary assessment process, the system lessens the workload for doctors, allowing them to focus more on providing high-quality care. Integration of real-time food data enhances healthcare systems' capacity for making decisions, which benefits mother outcomes and care.

The Pregnancy Women's Food Recommendation System utilizes AI to provide personalized, fast, and flexible nutrition recommendations, thereby promoting healthy pregnancies and benefiting mothers, healthcare providers, and the public.

## 6.2 Impact on Society & Environment

The Pregnancy Women's Food Recommendation System is expected to bring about several benefits to society. Environmental benefits of the pregnant women's meal suggestion system include the possibility of more effective and sustainable food management techniques. Software that can accurately detect 18 different classes of food items and provide specific dietary recommendations allows pregnant mothers to choose or capture images of the foods they should eat for dietary requirements. Food waste is prevented from rotting and costs are decreased as a result. Further reduction in food waste occurs when portions are tightly controlled to guarantee that food is consumed in sensible proportions for pregnant women. Foods that are seasonal, regionally sourced, and vegetarian can be preferred by the system if they are environmentally sustainable. By promoting the use of certain items, the system contributes to decreasing the environmental impact related to food production, preservation, and transport. Through the use of digital dietary monitoring, fewer physical resources are needed. The software allows pregnant women to track what they need to eat and get suggestions, doing away with the need for physical food diaries and printed dietary instructions. The transition to digital documentation helps reduce waste and the use of paper. This system can contribute to a more effective way for pregnant women to use healthcare resources by simplifying the nutritional management procedures. Healthcare facilities can save energy and money by reducing the number of in-person appointments for dietary recommendations. Also, if compared to conventional techniques, the application of AI and automated systems can reduce the overall energy requirements of food supply delivery. Dietary planning with AI and contemporary technology helps achieve more environmental goals by reducing worker costs and facilitating more sustainable operations. Since more patients might be accepted using the same or fewer resources, this can support the preservation of a strong economy.

There exists a beneficial relationship between the system's ability to decrease the frequency of patient visits to medical facilities for nutritional consults and the reduction in greenhouse gasses connected to transport. This reduces the overall environmental effect on medical services related to nutrition management during pregnancy.

## 6.3 Ethical Aspects

The Pregnancy Women's Food Recommendation System raises several ethical issues that need to be resolved to guarantee that it is implemented correctly. The method offers essential guidelines for ensuring that goods are delivered accurately, including food.
A woman's risk of health problems can increase significantly if she receives incorrect dietary identification or guidance throughout her pregnancy, particularly if she already has a medical condition like gestational diabetes.
Even with its high accuracy, the system still requires further real-world testing before it can be widely used; otherwise, it wouldn't be as trustworthy and beneficial as it already is. Guidelines to ensure mother's safety and procedures to reduce pregnant women's health conditions including failure. If the training does not involve the assembly of different diets and food items from different cultural backgrounds, there is a risk that the participant summarizes data.

To ensure fair and relevant recommendations, food databases covering a broad range of eating habits must be maintained. All users will receive timely and accurate nutritional information using routine evaluations of the system's performance across a wide range of demographic groups, which helps identify and correct biases with confidence and clarity.

Applying dietary suggestions to pregnant mothers appropriately requires the consideration of these ethical considerations. Maintaining user trust and ethical integrity while enhancing the system's positives requires ensuring accuracy, fairness, transparency, accountability, and support for affected individuals.

### 6.4 Sustainability Plan

For this system to evaluate pregnant women's food recommendation system from proof of concept to regular, widespread use partnerships and cooperative efforts, a special sustainability plan and commercial model are required. Working collaboratively, Dietitians, Software Programmers, and medical experts must work together to assess and enhance the system moving forward based on real-world usage.

They can increase the flexibility and consistency of the current techniques used in the healthcare sector with their innovative technology and specialized knowledge.

Collaborating with dietary and nutritional databases is important for ensuring that system recommendations are updated and suitable. One way to analyze the various options for distribution is to compare on-site installs with cloud-based Software-as-a-Service (SaaS) transmission. Every solution includes options for beginning costs, maintenance, security, flexibility of changes, and the ability to centralize ongoing model training. The Most affordable and effective method of action will need to be chosen by balancing all of these compromises.

It is necessary to look at a variety of funding possibilities to support the development and implementation of the system. You have three options: you can launch your own company on your own, obtain venture funding, or license to reliable health IT businesses.

It could be required to submit project and research funding applications to support more testing and development. As part of their go-to-market strategy, small clinics, and large medical facilities should focus on a wide range of healthcare providers.

Through being integrated into regular medical procedures, the Pregnancy Women's Food Recommendation System can give pregnant mothers exact and consistent nutritional guidance. Innovative technology not only alone guarantees the system's sustainability and scalability; a network of partners, finance, and operational protocols must be formed.

## 6.5 Summary

This section dealt with the various impacts of the Pregnancy Women's Food Recommendation System, highlighting its potential benefits, environmental implications, ethical quandaries, and sustainable methodologies.

First of all, the system has a big social impact since it wants to improve the health of both the mother and the newborn baby by giving pregnant women, especially those with diabetes, individualized dietary suggestions. The system improves pregnant mother's ability to make informed meal choices by utilizing cutting-edge technology such as YOLOv8 to provide precise food identification and prompt dietary advice.

The method encourages the use of seasonal, locally sourced, and vegetarian foods while minimizing food waste, all of which contribute to sustainable food management practices. By replacing paper-based nutritional tracking with digital technology, fewer physical resources are needed, which helps preserve the environment. The sustainability plan also highlights the necessity of software developers, dietitians, and other medical professionals working together to improve and optimize the system in light of actual user needs. Cost, maintenance, security, and scalability must all be balanced, which requires investigating various distribution options like cloud-based SaaS and on-site installations.

Improving the health of pregnant women and their babies’ results, addressing ethical concerns, and promoting sustainable practices are all possible with the Pregnancy Women's Food Recommendation System. With the right cooperation and support, this system can become a crucial resource for healthcare during pregnancy, benefiting pregnant women, healthcare providers, and the community in general.

# CHAPTER 7
# CONCLUSION AND FUTURE WORK

## 7.1 Conclusions

In conclusion, we have developed an innovative real-time food detection and recommendation system specifically for pregnant women with diabetes. By utilizing the YOLOv8 model, the system achieved high detection accuracy with a 0.882 mAP50 score in detecting and classifying 18 different food categories from images. It allows the users to receive personalized dietary recommendations. The integration of this model into a mobile application (mDiet), improved with TensorFlow Lite for efficient on-device inference that presents its practical applicability and effectiveness in real-world scenarios. The mobile app (mDiet), combines food detection from images and a recommendation system connected to Firebase that provides users with dietary suggestions based on detected food items and personalized health parameters. This approach significantly improves dietary management for pregnant women, helping them make informed food choices to better regulate blood glucose levels and overall health.

Overall, the project emphasizes the potential of leveraging advanced computer vision techniques like YOLOv8 to handle specific dietary needs. It showcases a practical application that not only meets health requirements but also improves user experience through accurate food detection and relevant recommendations.

## 7.2 Further Suggested Works

In our current project, we have got the expected result from our model and recommendation system. But over time this project can develop and increase its robustness. In future phases of this project, we few plans to expand the system,

- We want to add more food classes specifically related to the dietary needs of pregnant women with diabetes.
- We also have a plan to work with a larger dataset that would contain more images with different context and lighting conditions than our current dataset.
- We also intend to explore different models to improve detection accuracy beyond the current YOLOv8 model, ensuring the system remains robust and reliable.

- More importantly in our project, the recommendation system is one of the crucial parts. As we have just initiated the starting phase of this system in our project, currently based on predefined suggestions, will be further developed to become more automated and dynamic, can generate personalized advice based on real-time data, and user feedback

By executing these enhancements, we aim to create a comprehensive and adaptive tool that significantly aids in dietary management for pregnant women with diabetes.

### 7.3 Limitations/ Conflict of Interests

The study faced limitations related to the specificity of the dataset, including a small sample size and a limited number of food classes, which may not fully address individual dietary restrictions. Additionally, the predefined parameters within the recommendation system also limit its adaptability to diverse dietary needs. Cultural and economic factors also play a role in dietary choices, presenting challenges in creating a universally applicable system. There is a conflict of interest in this project because we are working with a sensitive domain which is pregnant women with diabetes. Based on the food detection and user health information advice from the doctor is always acceptable. However, an automated system may face trust and correctness issues so there is a conflict when it will be used in real life.

# REFERENCES


[1] Mak, J. K., Pham, N. M., Lee, A. H., Tang, L., Pan, X. F., Binns, C. W., & Sun, X. (2018). Dietary patterns during pregnancy and risk of gestational diabetes: a prospective cohort study in Western China. Nutrition Journal, 17(1), 1-11.

[2] Ravelli, M. N., & Schoeller, D. A. (2020). Traditional self-reported dietary instruments are prone to inaccuracies and new approaches are needed—frontiers in Nutrition, 7, 90. doi:10.3389/fnut.2020.00090

[3] Elmadfa, I., & Meyer, A. L. (2014). Developing suitable methods of nutritional status assessment: a continuous challenge. Advances in Nutrition (Bethesda, Md.), 5(5), 590S-598S. doi:10.3945/an.113.005330

[4] Nitish Nag, Aditya Narendra Rao, Akash Kulhalli, 2019. Flavor-enhanced food recommendations. arXiv:1904.05331v2 [cs.SI] 21 Aug 2019.

[5] Jiang L, Qiu B, Liu X, Huang C, Lin K. Deepfood: food image analysis and dietary assessment via deep model. IEEE Access. 2020;8:47477–89.

[6] Yang, S., Chen, M., Pomerleau, D., Sukthankar, R.: Food recognition using statistics of pairwise local features. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2249–2256 (2010)

[7]. Rostami, M., Muhammad, U., Forouzandeh, S., Berahmand, K., Farrahi, V., & Oussalah, M. (2022). An effective explainable food recommendation using deep image clustering and community detection. Intelligent Systems with Applications, 16, 200157.

[8]. Romadhon, F., Rahutomo, F., Hariyono, J., Sutrisno, S., Sulistyo, M. E., Ibrahim, M. H., & Pramono, S. (2023). Food Image Detection System and Calorie Content Estimation Using Yolo to Control Calorie Intake in the Body. In E3S Web of Conferences (Vol. 465, p. 02057). EDP Sciences.

[9]:Dewantara, B. S. B., Devy, A. Z., & Bachtiar, M. M. (2021, September). Recognition of Food Material and Measurement of Quality using YOLO and WLD-SVM. In 2021 International Electronics Symposium (IES) (pp. 545-551). IEEE.

[10].Nadeem, M., Shen, H., Choy, L., & Barakat, J. M. H. (2023). Smart diet diary: real-time mobile application for food recognition. Applied System Innovation, 6(2), 53.

[11] Sun, J., Radecka, K., & Zilic, Z. (2019). Foodtracker: A real-time food detection mobile application by deep convolutional neural networks. arXiv preprint arXiv:1909.05994.

[12] Liu, Y. C., Onthoni, D. D., Mohapatra, S., Irianti, D., & Sahoo, P. K. (2022). Deep-learning-assisted multi-dish food recognition application for dietary intake reporting. Electronics, 11(10), 1626.

[13] Krizhevsky, I Sutskever, and G E Hinton, “Imagenet classification with deep convolutional neural networks.,” In Advances in Neural Information Processing Systems, 2012.

[14] Anthimopoulos M.M., Gianola L., Scarnato L., Diem P., Mougiakakou S.G. A food recognition system for diabetic patients based on an optimized bag-of-features model. IEEE J. Biomed. Health Inform. 2014;18:1261–1271. doi: 10.1109/JBHI.2014.2308928.

[15] Chen M.-Y., Yang Y.-H., Ho C.-J., Wang S.-H., Liu S.-M., Chang E., Yeh C.-H., Ouhyoung M. SIGGRAPH Asia 2012 Technical Briefs (SA'12) Association for Computing Machinery; New York, NY, USA: 2012. Automatic Chinese food identification and quantity estimation; pp. 1–4.

[16] Beijbom O., Joshi N., Morris D., Saponas S., Khullar S. Menu-Match: Restaurant-Specific Food Logging from Images; Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision; Waikoloa, HI, USA. 5–9 January 2015; pp. 844–851.

[17] Kawano Y., Yanai K. Food image recognition with deep convolutional features; Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp'14 Adjunct); Seattle, WA, USA.

[18] Hassannejad H., Matrella G., Ciampolini P., de Munari I., Mordonini M., Cagnoni S. Food Image Recognition Using Very Deep Convolutional Networks; Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management (MADiMa'16); Amsterdam, The Netherlands. 16 October 2016; New York, NY, USA: Association for Computing Machinery; 2016. pp. 41–49.

[19] X. Yin, Z. Yu, Z. Fei, W. Lv, and X. Gao, "PE-YOLO: Pyramid Enhancement Network for Dark Object Detection." arXiv, Jul. 20, 2023. doi: 10.48550/arXiv.2307.10953.

[20] D. Reis, J. Kupec, J. Hong, and A. Daoudi, "Real-Time Flying Object Detection with YOLOv8." arXiv, May 17, 2023. doi: 10.48550/arXiv.2305.09972.

[21] Emmanuel WRS, Minija SJ. Fuzzy clustering and whale-based neural network to food recognition and calorie estimation for daily dietary assessment. Sadhana 2018; 43(5). doi: 10.1007/s12046-018-0865-3

[22] J. Zhang, J. Zhang, K. Zhou, Y. Zhang, H. Chen, and X. Yan, "An Improved YOLOv5-Based Underwater Object-Detection Framework," *Sensors*, vol. 23, no. 7, Art. no. 7, Jan. 2023, doi: 10.3390/s23073693.

[23] Zhu F., Bosch M., Khanna N., Boushey C.J., Delp E.J. Multiple Hypotheses Image Segmentation and Classification With Application to Dietary Assessment. IEEE J. Biomed. Health Inform. 2015;19:377–388. doi: 10.1109/JBHI.2014.2304925

[24] Sahoo, D., Hao, W., Ke, S., Xiongwei, W., Le, H., Achananuparp, P., ... & Hoi, S. C. (2019, July). FoodAI: Food image recognition via deep learning for smart food logging. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2260-2268).

[25] Ultralytics, available at https://www.ultralytics.com/, last accessed on 12 July 2024.

[26] MakeSenseAi, available at https://www.makesense.ai/, last accessed on 01 May 2024.

# Appendix A

## Course Outcomes, Complex Engineering Problems (EP) and Complex Engineering Activities (EA) Addressing

**Title:**

**mDiet: A DIETARY MANAGEMENT SYSTEM OF DIABETES IN PREGNANCY USING MACHINE LEARNING**

**Student ID: 201-15-13976, 201-15-13746**

## CO Description for FYDP

| CO | CO Descriptions | PO |
|---|---|---|
| **Phase -I** | | |
| **CO1** | Integrate recently gained and previously acquired knowledge to identify a **object detection** problem for the Final Year Design Project (FYDP) | **PO1** |
| **CO2** | Analyze different aspects of the goals in designing a solution for this FYDP | **PO2** |
| **CO3** | Explore diverse problem domains through a literature review, delineate the issues, and establish this goals for the FYDP | **PO4** |
| **CO4** | Perform economic evaluation and cost estimation and employ suitable project management procedures throughout the development life cycle of the FYDP | **PO11** |
| **Phase -II** | | |
| **CO5** | Design and develop technical solutions and system components or processes that meet specified requirements, ensuring compliance with public health and safety standards, as well as considering cultural, socioeconomic, and environmental factors in this FYDP | **PO3** |
| **CO6** | Choose and apply appropriate methodologies, resources, and contemporary engineering and IT technologies to address complex engineering processes, encompassing prediction and modeling, while adhering to relevant constraints in this FYDP | **PO5** |
| **CO7** | Analyze societal, health, safety, legal, and cultural considerations, along with associated responsibilities, in the context of professional engineering practice and the resolution of this problem, employing logical reasoning guided by contextual understanding. | **PO6** |
| **CO8** | Comprehend and evaluate the enduring sustainability and impact of professional engineering endeavors in addressing intricate engineering challenges within social and environmental frameworks. | **PO7** |

| | | |
|---|---|---|
| **CO9** | Implement ethical principles and adhere to professional standards and norms in this FYDP | **PO8** |
| **CO10** | Capable of operating proficiently both individually and as a team member or leader across diverse teams and interdisciplinary settings in this FYDP. | **PO9** |
| **CO11** | Proficiently communicate with the engineering community and broader society regarding complex engineering endeavors, including the ability to comprehend and generate comprehensive reports and design documentation, as well as provide and receive clear instructions throughout this FYDP. | **PO10** |
| **CO12** | Acknowledge the importance of self-directed and life-long learning within the evolving landscape of technology, and possess the readiness and capability to engage in lifelong learning endeavors. | **PO12** |

## Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP), and Attainment of Complex Engineering Activities (EA)

### Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP):

| SN | EP Definition | Attainment | CO | Justification (with Knowledge Profile) | References |
|---|---|---|---|---|---|
| **1.** | **EP1: Depth of Knowledge required** | Yes | CO1, CO2, CO3, CO5, CO6, CO7 and CO8 | This machine learning project demonstrates **fundamental engineering (K3)** principles by employing machine learning techniques using YOLO, data augmentation techniques, and diverse YOLO architectures for image processing and object detection tasks. The project demonstrates **specialist knowledge (K4)** by conducting transfer learning with advanced CNN architectures like YOLOv5 and YOLOv8, enhancing food detection accuracy in food images, crucial for computer-aided diagnosis. | **Page no:** [15-21] **Section:** [3.2.1.3, 3.2.4.1 ] **Page no:** [20-21] **Section:** [3.2.4.2] |
| | | | | The project applies **engineering practice & design (K5)** by developing an mobile | **Page no:** [35-36] |

| | | | | | |
|---|---|---|---|---|---|
| | | | | application to experiment the model we built. The project addresses **engineering practice & technology (K6)** by employing YOLOv5 and YOLOv8 models. | **Section:** [5.2.2] **Page no:** [24-27] **Section:** [4.2] |
| | | | | This project ensures to **K8 (Research Literature)** by synthesizing insights from recent studies, to advance object detection using deep learning techniques, showcasing a comprehensive understanding of current methodologies. | **Page no:** [6-12] **Section:** [2.2, 2.3] |
| **2.** | **EP2: Range of Conflicting Requirements** | Yes | CO2, and CO7 | This project addresses **EP-2** by recognizing the hurdles in food detection, including limitations of traditional methods and the complexities of integrating machine learning, transfer learning and deep CNNs. Through comparative analysis, it confronts challenges in understanding spatial distributions, offering insights for refining diagnostic methodologies. | **Page no:** [9-12] **Section:** [2.3, 2.4] |
| **3.** | **EP3: Depth of analysis required** | Yes | CO2, and CO6 | This project addresses **EP-3** by meticulously comparing experimental outcomes, highlighting Transfer Learning as the chosen significant solution for enhancing food detection amidst multiple potential approaches. | **Page no:** [30-34] **Section:** [5.2.1] |

| | | | | | |
|---|---|---|---|---|---|
| **4.** | **EP4: Familiarity of Issues** | Yes | CO8 | This project's interdisciplinary approach extends beyond computer science and engineering, impacting medical diagnostics in respiratory issues like dietary management in pregnancy with diabetes, contributing to advancements in public health and healthcare practices which indicates **EP-4**. | **Page no:** [2-3] **Section:** [1.3, 1.5] |
| **5.** | **EP5: Extends of application codes** | No | CO5 | N/A | N/A |
| **6.** | **EP6: Extends of stakeholders involved and conflicting** | No | CO8 | N/A | N/A |
| **7.** | **EP7: Interdependence** | Yes | CO5 | This project's comprehensive approach addresses high-level problems by integrating various components across data collection, statistical data analysis, and proposed methodology, ensuring a holistic solution to complex challenges in medical diagnostics which ensures **EP-7**. | **Page no:** [13-19] **Section:** [3.2, 3.2.1, 3.2.1.1, 3.2.3, 3.2.4] |

**Addressing CO11 with Complex Engineering Activities (EA) [Some or all of the following]:**

| SN | EA Definition | Attainment | CO | Justification | References |
|---|---|---|---|---|---|
| **1.** | **EA1: Range of resources** | Yes | CO11 | Our project utilizes diverse resources such as high-performance computing infrastructure, GPUs, machine/deep learning frameworks, annotated datasets, and ethical considerations to ensure systematic research and contribute to advancements in object detection through machine learning. | **Page no:** [15-23] **Section:** [3.2.1.2, 3.2.4.1, 3.3] |
| **2.** | **EA2: Level of interaction** | No | | N/A | N/A |

| | | | | | |
|---|---|---|---|---|---|
| **3.** | **EA3: Innovation** | No | | N/A | N/A |
| **4.** | **EA4: Consequences for society and the environment** | Yes | | This project contributes to society by improving healthcare through advanced food detection methods and building recommendation system for pregnant woman, while also promoting environmental sustainability by employing efficient computational resources and adhering to ethical guidelines for patient data privacy. | **Page no:** [40-42] **Section:** [6.1, 6.2, 6.4] |
| **5.** | **EA-5: Familiarity** | Yes | | This project expands upon existing research by examining a novel approach in object detection through YOLO models and deep CNNs, demonstrated through preliminary terminologies and a comprehensive comparative analysis, offering new insights into the field. | **Page no:** [9-12] **Section:** [2.3] |

**Addressing CO (4, 9, 10, and 12):**

| SN | COs | Attainment | Justification | References |
|---|---|---|---|---|
| **1** | CO4 | Yes | This project addresses **CO4** by integrating effective project management and financial oversight, ensuring meticulous planning, resource allocation, and budget estimation for optimal resource utilization throughout the research lifecycle. | **Page no:** [22-23] **Section:** [3.4] |
| **2** | CO9 | Yes | The project demonstrates adherence to ethical principles by prioritizing patient privacy, obtaining informed consent, | **Page no:** [41] **Section:** |

| | | | | |
|---|---|---|---|---|
| | | | and transparently documenting the research process, ensuring responsible knowledge dissemination and societal well-being through the ethical application of advanced healthcare technologies which comply with **CO9**. | [6.3] |
| **3** | CO10 | No | N/A | N/A |
| **4** | CO12 | Yes | The project's dedication to continuous learning **(CO12)** and adaptation within the dynamic technological landscape is reflected in its comprehensive data collection, rigorous statistical analysis, meticulous methodology development, and thorough experimental results and analysis, showcasing a commitment to staying updated and refining techniques to address modern challenges. | **Page no:** [13-14, 30-34, 44-45] **Section:** [3.2, 3.3, 3.2.1.1, 5.2.1, 5.2.2, 5.3] |

# mDiet: A DIETARY MANAGEMENT SYSTEM OF DIABETES IN PREGNANCY USING MACHINE LEARNING

ORIGINALITY REPORT

| 18% | 12% | 8% | 11% |
| --- | --- | --- | --- |
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
| --- | --- | --- |
| 1 | Submitted to Daffodil International University<br>Student Paper | 5% |
| 2 | dspace.daffodilvarsity.edu.bd:8080<br>Internet Source | 2% |
| 3 | www.researchgate.net<br>Internet Source | 2% |
| 4 | Submitted to Liverpool John Moores University<br>Student Paper | 1% |
| 5 | arxiv.org<br>Internet Source | <1% |
| 6 | chped.net<br>Internet Source | <1% |
| 7 | Maryam Abbasi, Filipe Cardoso, Pedro Martins. "Chapter 31 Comparison ofImage Processing andClassification Methods foraBetter Diet Decision-Making", Springer Science and Business Media LLC, 2023<br>Publication | <1% |