**Project Proposal:**

*Fruit Classifier and Freshness Evaluator Model*

Contributors:

Megan Gross - 018230836

Lindsey Raven - 018200819

Sonal Prasad - 017475432

Ashley Irawan - 015902406

Pushpal Patil - 013166127

Peter Conant - 012155624

[**1. Project Concept 3**](#_giz0x17pwwgq)

[1.1 Project Abstract 3](#_ku0fimh3xv6j)

[1.2 Project Background 3](#_9twl3fyh9rtc)

[1.2 Preliminary Model Architecture 4](#_xjprawrhd0dk)

[1.2.1 Preliminary Plan 4](#_2oxd6z297wg3)

[1.2.2 Extended Plan 4](#_knostzxrbfro)

[1.2.3 Datasets 5](#_5nsddxe4cgg6)

[1.2.4 Learning Models 5](#_6tr8paw717jb)

[1.3 Goals and Objectives 6](#_994xgntvuqg3)

[**2. Project Management 7**](#_bm0gcfld4rwv)

[2.1 Member Short Bio/Introduction 7](#_spyj13yf6hfk)

[2.1.1 Sonal Prasad 7](#_5kbpckpfbmw2)

[2.1.2 Peter Conant 8](#_889zvbc3z8t2)

[2.1.3 Megan Gross 8](#_c9gcnj60v0oz)

[2.1.4 Lindsey Raven 8](#_wa128suuy07q)

[2.1.5 Ashley Irawan 8](#_e73599mk7h2c)

[2.1.6 Pushpal Patil 8](#_hpv504nks7no)

[2.2 Task Distribution 9](#_o1hmor4wx3a5)

[2.2.1 Gantt Chart 10](#_kean27bcno7w)

[2.2.2 Project Milestones 11](#_ydxbtn4yljyi)

[2.2.3 Group Communication Tools & Techniques 12](#_9mm9kku8h3e9)

[2.2.4 Project Tools/Frameworks 12](#_b5yzcsaxn7ok)

[2.2.4.1 Learning Framework 12](#_rne4yz43buyi)

[2.2.4.2 Versioning Software 12](#_hdm30epb8bxx)

[**References 12**](#_if19omdlfkgd)

# **1. Project Concept**

## **1.1 Project Abstract**

*Grocery retailers tend to not capitalize on redistribution of expiring fruit to food banks to maximize profit. This leads to excessive food waste. Proposed in the following sections is a project proposal detailing the creation of a model that, given an image, can classify the fruit contained within and assess how close the fruit is to expiration. The models that will be used are ResNet, Vision Transformers, and YOLOv8. They will be analyzed by performance metrics such as accuracy, precision, recall, and ROC curve.*

## **1.2 Project Background**

Fruit is a key diet component for attaining a healthy lifestyle and life longevity. As a result, there is a high demand for the global fruit market. In the US alone, the fruits and vegetables market size was valued at 96.26 billion in 2023 [[1]](#sl9biqyp0b4v).

This demand is high despite the fact that fruit as a product is quite volatile: it has a shelf life that is short. Grocery stores attempt to get the product into consumer hands before its expiry, yet a large percentage of product still becomes waste. In 2022, it was estimated that grocery stores had 5 million tons of surplus food, 30% of which went straight to landfill [[2]](#zb04nf7b4p9z).

Since a percentage of fruits shelved are destined for expiration, a solution is to give the food to those in need at no cost. However, grocery stores will only pull products from the shelves for donations at the last possible moment. As a result, grocery stores tend to miss the window of opportunity, and the food inevitably goes to waste.

It is the intention of this proposal to attempt to solve this problem: to alert grocery retailers when fruits, which have a short shelf life, are approaching their expiration and are in the best interest of the retailer to alert institutions to redistribute the food for the benefit of others. In this proposal, we present an outline for a project that, given images of fruit, will identify their species, and correctly identify when they are close to expiration.

## **1.2 Preliminary Model Architecture**

### 1.2.1 Preliminary Plan

As stated, the goal of this project is to produce a model that, given an image, can identify fruit and quantify how close it is to expiring. The model can be broken down into achieving three main tasks:

1) The identification of the fruit within the image.

2) The classifying of the fruit

3) The quantification of how close the fruit is to expiring.

There are then several approaches that can be taken to address these steps:

(A) A single machine learning model could be used to perform (1-3).

(B) A model could be used to object identify (1) and then another to classify (2-3).

(C) A model could be used for each step, having the output of one feed into the next.

As a preliminary approach, it is the proposal of this document to pursue (A). Breaking down the model with steps (2) and (3) may help resolve issues related to overfitting or underfitting, but by pursuing approach (A), this project can focus on classifying more variations of fruit instead of the model breakdown.

As an attainable goal, while maintaining learnings: the contributors of this project will split into three groups of three and focus on beginning with using existing models and adapting them to achieve (A) with the most basic case: classifying a single fruit and determining whether or not it is rotten.

### 1.2.2 Extended Plan

Depending on the success of the model (discussed in [1.3](#u63w9cb9knex)) and the expected timeframe (discussed in [2.2](#kznc7xqa2twk)) the models can then be adapted to:

1. Add more fruits to the model classifications
2. Quantify the ripeness scale
3. Improve the performance metrics by changing the architecture of the existing models

In this case: 1) - 3) are stretch goals of the project and are outside the scope of phase 1. Should timeframe or results allow: the beginning of these tasks will be added to phase 1.

### 1.2.3 Datasets

We will be utilizing several datasets from Kaggle. The first is of ripe and rotten apples, oranges, and bananas, where each category (i.e. ripe apples, rotten oranges, etc.) from the training data has 1400+ images. The source also provides a separate set of testing data with the same categories as the above [[3]](#6fz15faf0o73).

Our second dataset is “Fruit and Vegetable Disease (Healthy vs Rotten)” which has 29 thousand files of 14 types of fruits and vegetables. These include apples, bell peppers, and guava. Each has their own fresh and rotten image set. All images have a mixture of flat white backgrounds and busy backgrounds [[4]](#xfsy00p6tty).

When training the model for object detection and classification we will use “Fruits by YOLO”. This dataset has nearly 3000 images with labels for the location and type of fruit in the image. Some images have multiple fruits and types of fruits within them, and ideally this will increase the robustness of our model [[5]](#fotxpmtwdzr1).

### 1.2.4 Learning Models

Convolutional Neural Networks are responsible for many breakthroughs in image classification in the early 2010s [[6]](#obkk7fwm5j5z) and have previously shown high accuracy for classifying fresh and rotten fruits [[7]](#8b1vtvrr346r). One of the models we will be using is the ResNet50 model which is a mid-sized CNN. Each convolutional layer uses a residual block to compare its input to its output and bottleneck architecture to reduce and restore dimensionality within the layer. We expect this model to perform very well in our project to use as a baseline.

The Vision Transformer model (ViT) is a transformer model that focuses on computer vision tasks. ViT has been fairly consistently outperforming Convolutional Neural Networks (CNN) and was first introduced as an optimal CNN competitor for image classification in 2021 at the International Conference on Learning Representations [[8]](#9zt2pojmivja). Although ViT’s training time is less than that of CNNs, ViT requires significantly more training data than CNNs. While ViT has been found to do well in image classification [[9]](#eavvvq7h1k2f), we are uncertain how effective it is for our use case of classifying fruit as ripe or rotten. Although accuracy tends to be high for ViT, we have a small dataset of approximately 1500 images per fruit for each category (i.e. ripe, rotten). With further experimentation, such as data augmentation or pre-training, we may be able to reach an increased accuracy compared to the standard CNN model.

A common issue with regular object detection models is detecting objects that are overlapping with one another. One way to counteract this is to use instance segmentation. YOLOv8-seg segmentation model is a variant of the YOLOv8 object detection model that performs real-time instance segmentation.[[10]](#bex3wthqhn11) YOLOv8 models are built on a CSP-DarkNet53. [[13]](#vmbliwg86f49) When a CSPNet is applied to CNN backbones like ResNet, it is not only able to reduce computation cost and memory usage, but also outperforms plain ResNet on inference speed and accuracy when tested on the ImageNet dataset. Since YOLOv8 is often mentioned in comparison to ResNet, we have selected this as our last model for analysis and application to fruit classification. It also has been used for tomato classification before [[11]](#o8j06j9ttpuz). We will be comparing results of this model to our other models, and determining whether the inherent differences in architecture as well as activation functions results in a significant difference in results. Of note: if we encounter any issues with YOLOv8-seg or find it is not a good fit, we plan to switch models to research DETR [[12]](#gar3gfib9ebb).

## **1.3 Goals and Objectives**

The preliminary goal of this project is to research 3 existing models, and possibly improve their design, to detect and classify one fruit type. In order to properly report and quantify the success of our models: we will be using standard success metrics for classification learning models. These metrics include:

1. **Accuracy** - Total successful model guesses (eg. ripe vs rotten) over all guesses
2. **Precision** - Total correct positive guesses over all model positive guesses (eg. all correct positive ripe guesses over all positive guesses)
3. **Recall** - Total positive correct guesses over all positive samples (eg. all correct positive ripe guesses over all positive ripe samples)
4. **Receiver Operating Characteristic (ROC)** - graph that shows the true positive rate over the false positive rate given differing thresholds for classification

Every intermediary model will be reporting / visualizing these metrics in order to report progress. As a preliminary goal for our final model, we would like to achieve the following:

1. Accuracy, Precision, Recall - 85% accuracy in reporting the ripeness quality of a given fruit and its fruit type classification
2. ROC Curve - No specific metric goal, but would like to target a very high TRP with very low threshold for both fruit classification and ripeness classification

We will use a versioning system (see [2.2.4.2](#8lxn7hgzfk4)) which will document each significant model change we make, and corresponding results.

# **2. Project Management**

## **2.1 Member Short Bio/Introduction**

### 2.1.1 Sonal Prasad

Sonal Prasad is currently a Master’s student at San Jose State University, majoring in Artificial Intelligence, and previously completed her B.S. Computer Science degree from UC San Diego. Sonal has 3 years of work experience as a software engineer, including working full stack and on an AI infrastructure team.

### 2.1.2 Peter Conant

Peter Conant is in the Master in Artificial Intelligence program at SJSU. He completed his undergraduate at SDSU and worked as a research assistant applying CNNs to marine mammal hydroacoustics. He plans to make a career in autonomous systems and is very interested in the search and rescue research of Professor Jun Liu.

### 2.1.3 Megan Gross

Megan Gross is in the Master in Artificial Intelligence program at CSU San Jose. She obtained her Bachelors in computer science at CSU Sacramento. She has been interested in AI since studying abroad in Australia at the University of Technology, Sydney in fall 2022 where she took an introductory course and an image processing course. She is interested in utilizing AI for accessibility purposes to make the world a more equitable and accessible place for all.

### 2.1.4 Lindsey Raven

Lindsey Raven is a professional engineer who has worked in the firmware industry with enterprise storage devices for the last half decade. She has achieved a bachelor's degree in Electrical Engineering, and is currently pursuing a masters in Artificial Intelligence at San Jose State university. Her prior research involves improving feature detection and description algorithms for image processing.

### 2.1.5 Ashley Irawan

Ashley Irawan is a Master’s student at San Jose State specializing in Artificial Intelligence. As a CS student at SJSU, she worked in building AI applications with LLMs and worked as a research assistant with Coral Vision Research group at San Jose State’s College of Science. A future entrepreneur, Ashley is interested in using AI to solve real world problems and aspires to create her own startup.

### 2.1.6 Pushpal Patil

Pushpal Patil is a 2nd year Master’s student at SJSU studying M.S. Software Engineering. She did her B.S. Software Engineering at SJSU as well. During her undergrad, she developed a passion for mobile app development and wants to learn more about integrating ML in her mobile app projects to provide a better user experience. She is interested in how ML/AI can be used to improve quality of life and wants to work towards building a product/company she is passionate about - be it her own or someone else’s.

## **2.2 Task Distribution**

The initial tasks of this project are training the Resnet50, Vision Transformer, and Segmentation models. We split these models evenly among the team members working collaboratively in groups of two: Megan and Sonal will create the Vision Transformer(ViT - B), Peter and Pushpal will create the Deep CNN(ResNet50 ), and Lindsey and Ashley will create the Segmentation Model (YOLOv8 / DETR). Each duo will be responsible for the training and evaluation of the model on fruit detection, fruit type classification, and ripe vs rotten classification.

### 2.2.1 Gantt Chart

Note that the [Placeholder Model 1][Fruit 1] work will be shared by 2 individuals and dates and timelines will be 3x as there will be 3 models being pursued by 3 sub groups. For details on extended plans and which models will be pursued, see the prior sections 1.2.2 and 1.2.4.

### 

### 2.2.2 Project Milestones

| **Date** | **Title** | **Description** | **Goal Requirements** |
| --- | --- | --- | --- |
| Sep. 25th | Project Proposal Finish | Complete this document | NA |
| Oct. 7th | Model Launch | Have the model running on framework and complete min. 1 epoch of training. | Classification -  1 Fruit, Ripe / Rotten  Accuracy - no requirement |
| Oct. 8th - 11th | Midterm Presentation | Present initial results | NA |
| Oct. 17th | Performance Check In 1 | Assess current results of existing models | Classification -  1 Fruit, Ripe / Rotten  Accuracy - 85% |
| Oct. 25th | Performance Check In 2 | Attempt to reach performance peak | Classification -  1 Fruit, Ripe / Rotten  Accuracy - 90% |
| Oct. 29th - 31st | Case Study Presentation | Research similar model | NA |
| Nov. 19th - 21st | Final Presentation | Final presentation dates | Classification -  1< Fruit, Ripe / Rotten Scale?  Accuracy - 95% |
| Nov. 19th | Final Report | Final Report due | Classification -  1< Fruit, Ripe / Rotten Scale?  Accuracy -95% |

### 2.2.3 Group Communication Tools & Techniques

The group is using Discord to communicate and coordinate. There will be weekly meetings on Mondays at 2PM. These meetings will be stand up meetings to reflect and address the previous week's task and what each group member completed. With this the team will be able to track progress and set goals for what to complete in the next week.

### 2.2.4 Project Tools/Frameworks

#### 2.2.4.1 Learning Framework

Several of our team members have some experience using tensorflow. To supplant this experience, we will be attempting to use the tensorflow framework to implement the models mentioned in [[1.2.4](#_6tr8paw717jb)]. Our current plan is to run our models locally. We may expand to running on the cloud if we find that we are limited by our hardware.

#### 2.2.4.2 Versioning Software

We will be using a versioning software, GitHub, to help document the various experiments and changes we make to each model. The structure of our repository will make it clear which experiments were performed, as well as the result findings. If it’s not clear from the architecture, the commits will also include details of what was changed / created for each model.

# References

[1] “U.S. fruit & Vegetables Market Size, Share | Industry Report, 2018-2025,” *Grandviewresearch.com*, 2018. <https://www.grandviewresearch.com/industry-analysis/us-fruit-vegetables-market>

[2] Bridget Reed Morawski, “How supermarkets and grocery stores contribute to food waste,” *one5c*, Mar. 14, 2024. <https://one5c.com/food-waste-supermarkets-136944209> (accessed Sep. 27, 2024).

[3] “Fruits fresh and rotten for classification,” *www.kaggle.com*. <https://www.kaggle.com/datasets/sriramr/fruits-fresh-and-rotten-for-classification> (accessed Sep. 27, 2024).

[4] Muhammad Subhan, 2024. “Fruit and Vegetable Disease (Healthy vs Rotten)” distributed by Kaggle, <https://www.kaggle.com/datasets/muhammad0subhan/fruit-and-vegetable-disease-healthy-vs-rotten> (accessed Sep. 27, 2024).

[5] Alexander Kapturov, 2024. “Fruits by YOLO - Fruits Detection” distributed by Kaggle <https://www.kaggle.com/datasets/kapturovalexander/fruits-by-yolo-fruits-detection> (accessed Sep. 27, 2024).

[6] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *arXiv.org*, Dec. 10, 2015. <https://arxiv.org/abs/1512.03385>

[7] “Fresh and Rotten Fruits Classification Using CNN and Transfer Learning | IIETA,” *Iieta.org*, 2023. <https://www.iieta.org/journals/ria/paper/10.18280/ria.340512> (accessed Sep. 27, 2024).

[8] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale”. *International Conference on Learning Representations*. <https://openreview.net/pdf?id=YicbFdNTTy>

[9] A. Berroukham, K. Housni and M. Lahraichi, "Vision Transformers: A Review of Architecture, Applications, and Future Directions," 2023 7th IEEE Congress on Information Science and Technology (CiSt), Agadir - Essaouira, Morocco, 2023, pp. 205-210, doi: 10.1109/CiSt56084.2023.10410015.

[10] Ultralytics, “YOLOv8,” *docs.ultralytics.com*, Nov. 12, 2023. <https://docs.ultralytics.com/models/yolov8/>

[11] G. Yang, J. Wang, Z. Nie, H. Yang, and S. Yu, “A Lightweight YOLOv8 Tomato Detection Algorithm Combining Feature Enhancement and Attention,” *Agronomy*, vol. 13, no. 7, pp. 1824–1824, Jul. 2023, doi: <https://doi.org/10.3390/agronomy13071824>.

[12] F. R. PhD, “DEtection TRansformer (DETR) vs. YOLO for object detection.,” *Medium*, Mar. 18, 2024. <https://medium.com/@faheemrustamy/detection-transformer-detr-vs-yolo-for-object-detection-baeb3c50bc3>

[13] O. M. Khare, S. Gandhi, Rahalkar, Aditya M, and S. Mane, “YOLOv8-Based Visual Detection of Road Hazards: Potholes, Sewer Covers, and Manholes,” *arXiv.org*, 2023. <https://arxiv.org/abs/2311.00073>

‌

‌

‌