PerishPredict: Smart Vision for Produce Spoilage Detection

Introduction

Our Team

Advaith Shankar

Andrew White

Sarah Stephens

Twinkle Panda

Varsha MJ

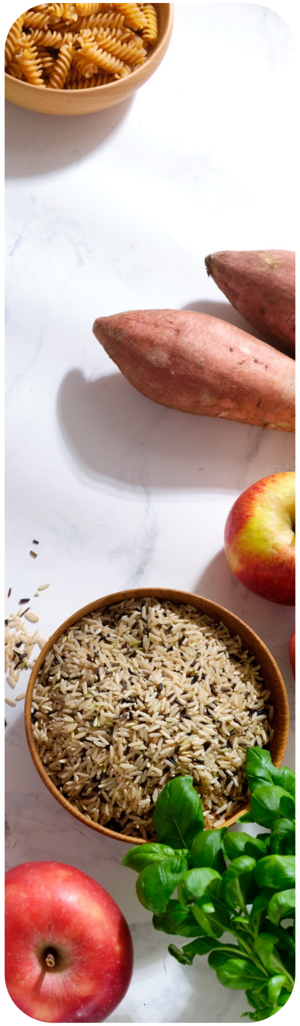
Agenda

1. Problem
2. Convolutional Neural Networks
3. Multimodal LLMs
4. Support Vector Machine
5. Contrastive Learning
6. Applications
7. Concluding Thoughts

Introduction

Problem

- Fresh produce is especially prone to damage or aging during its journey from farm to shelf
- Food waste is a significant issue in the grocery business, especially considering the low margins retailers make
- Food waste contributes to global emissions
- Spoiled produce can diminish customer satisfaction and reduce sales



Our Solution

Spoilage-Recognition Technology

By utilizing a pre-trained foundation model with image recognition capabilities, retailers can assess the quality of perishable products like fruits and vegetables before they are made available to consumers at stores, enhancing both profitability and customer experience.

About the Data

Dataset I

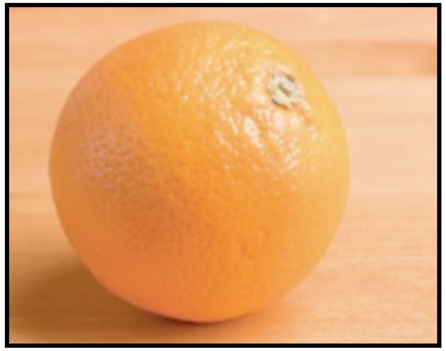
Approximately 13,000 images that include apples, oranges, and bananas, with both fresh and spoiled examples.

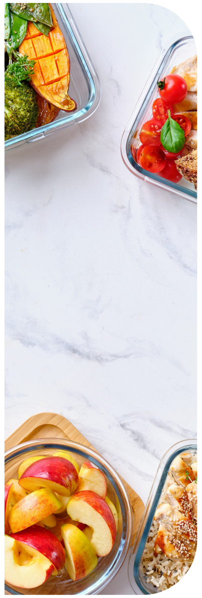
Dataset II

Approximately 30,000 images that include a wide variety of fruits and vegetables, with both fresh and spoiled examples.

Both datasets are from Kaggle

Each image is labeled either as "fresh" or "spoilt," providing a solid foundation for training a binary classification model.





Models

Convolutional Neural Networks

Definition

An artificial, feed-forward neural network that updates weights and biases through processes called backpropagation and gradient descent to determine an output.

Advantages

- Highly effective for image and video recognition due to:
  - Automatic Feature Extraction
  - Hierarchical Feature Learning
  - Translation Invariance
- Scalable to large datasets which allows them to solve highly complex problems

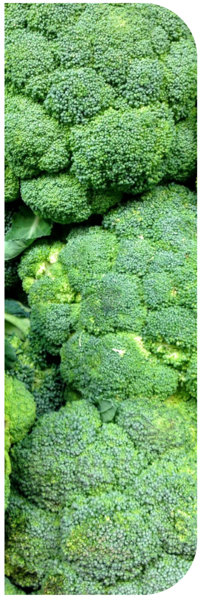
Disadvantages

- Depends on huge amounts of data
- High computational costs to train
- Can easily overfit (sometimes requires regularization to compensate)
- Low interpretability
- Hyperparameter tuning can be tricky

Types of Models

TensorFlow

(Tried ResNet-50 and EfficientNet, poor results)

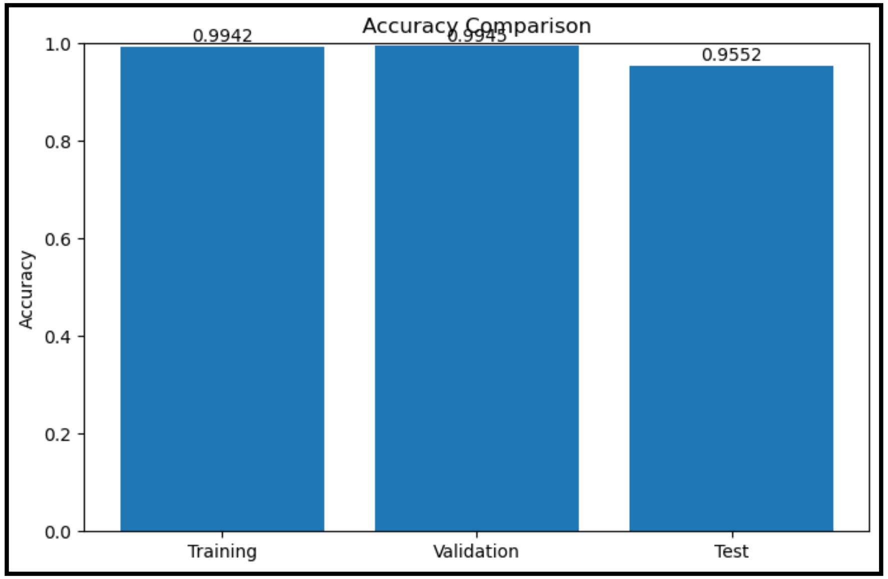


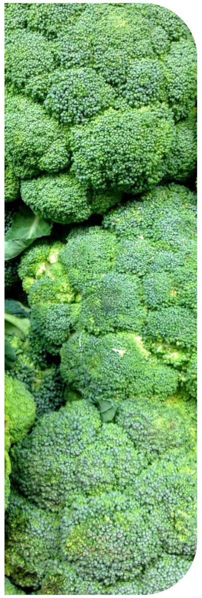
Models

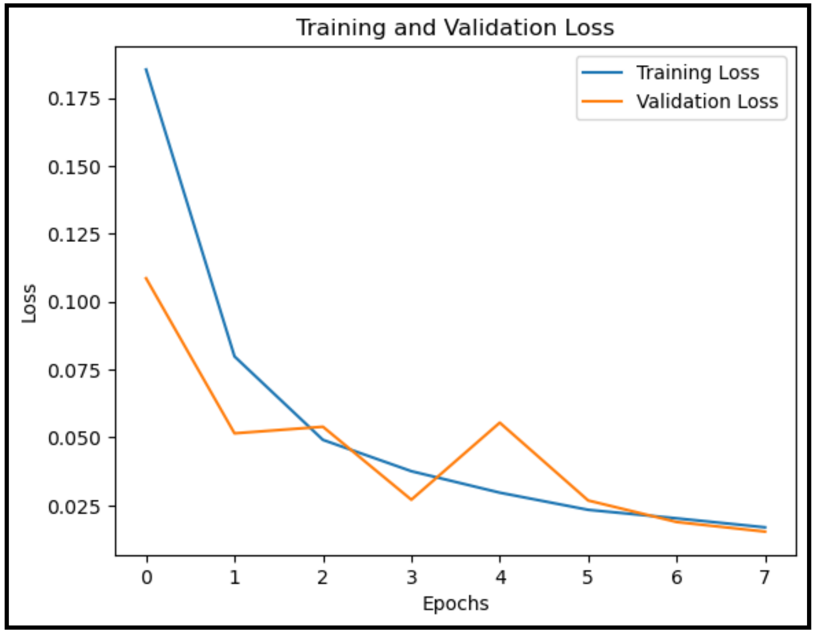
TensorFlow CNN

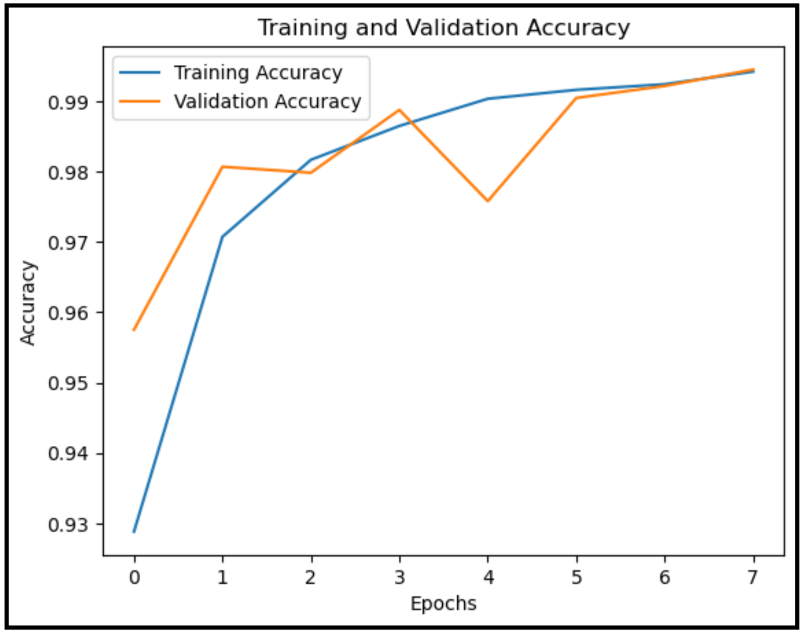
Comprehensive open-source ML framework

```
Training Accuracy: 0.9942
Validation Accuracy: 0.9945
Test Accuracy: 0.9552
```






Training and Validation Loss
Training Loss
Validation Loss
Loss
Epochs
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0
1
2
3
4
5
6
7


Training and Validation Accuracy
Training Accuracy
Validation Accuracy
Accuracy
Epochs
0.99
0.98
0.97
0.96
0.95
0.94
0.93
0
1
2
3
4
5
6
7

Models

Multimodal LLMs

Definition

AI systems (large language models) that can process multiple data type (modality) inputs to generate insights

Advantages

- Versatile—Can handle various “modalities”
- Wide range of applications
- Strong ability to capture context/nuance

Disadvantages

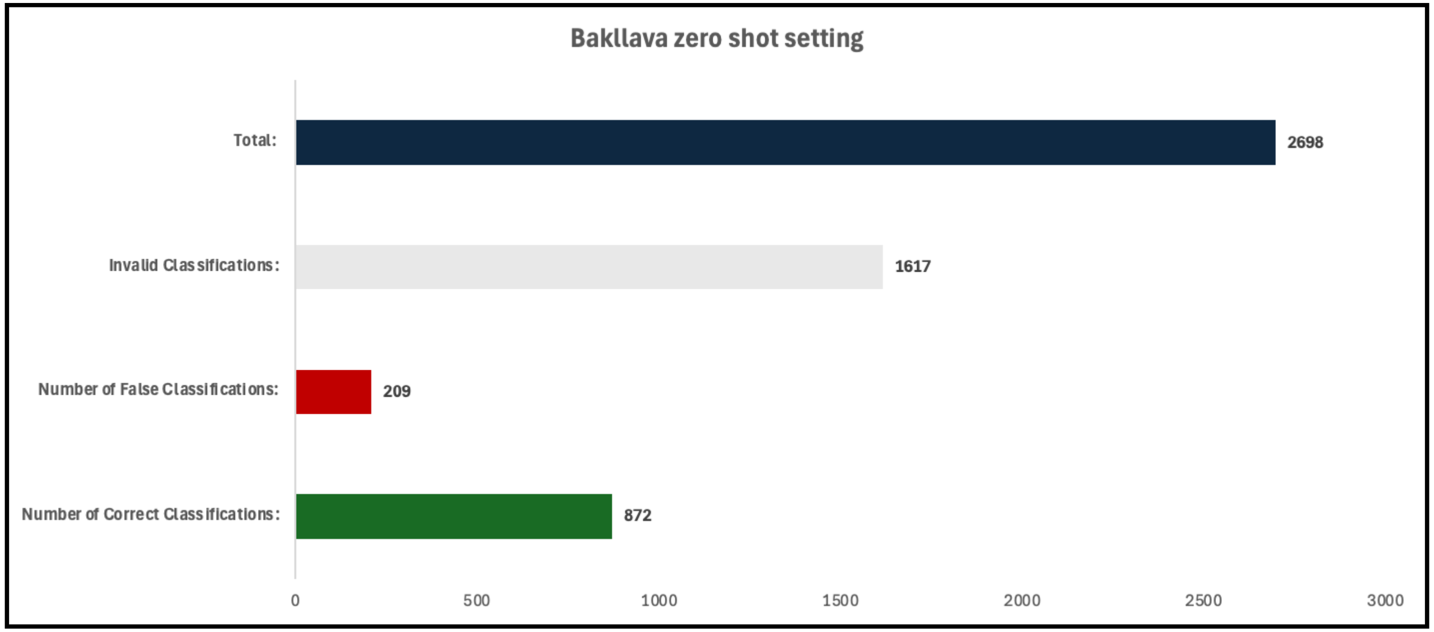
- Computationally demanding—complex architectures
- Relies on large datasets
- Hard to interpret

Types of Models

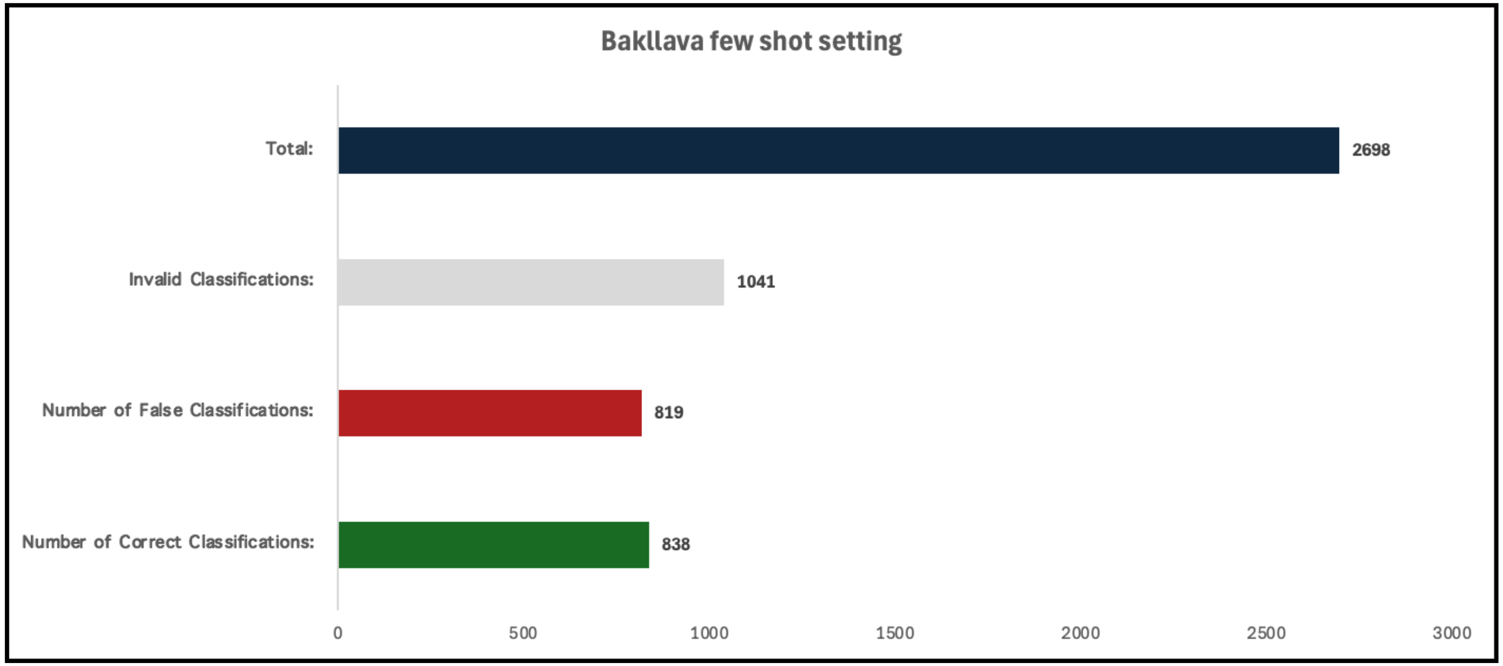
Llava-13b, BakLLaVA (Mistral 7B augmented with LLaVA architecture)

The models were set up on local using Ollama

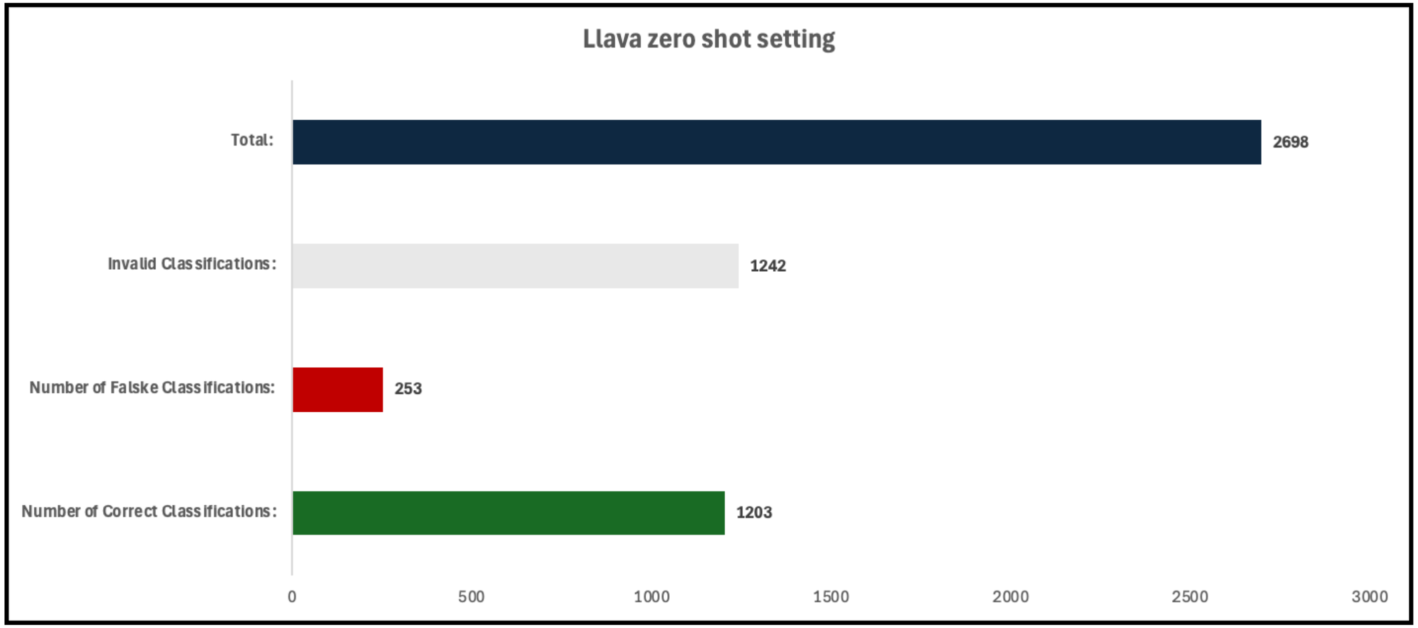
Zero Shot Prompting - Bakllava



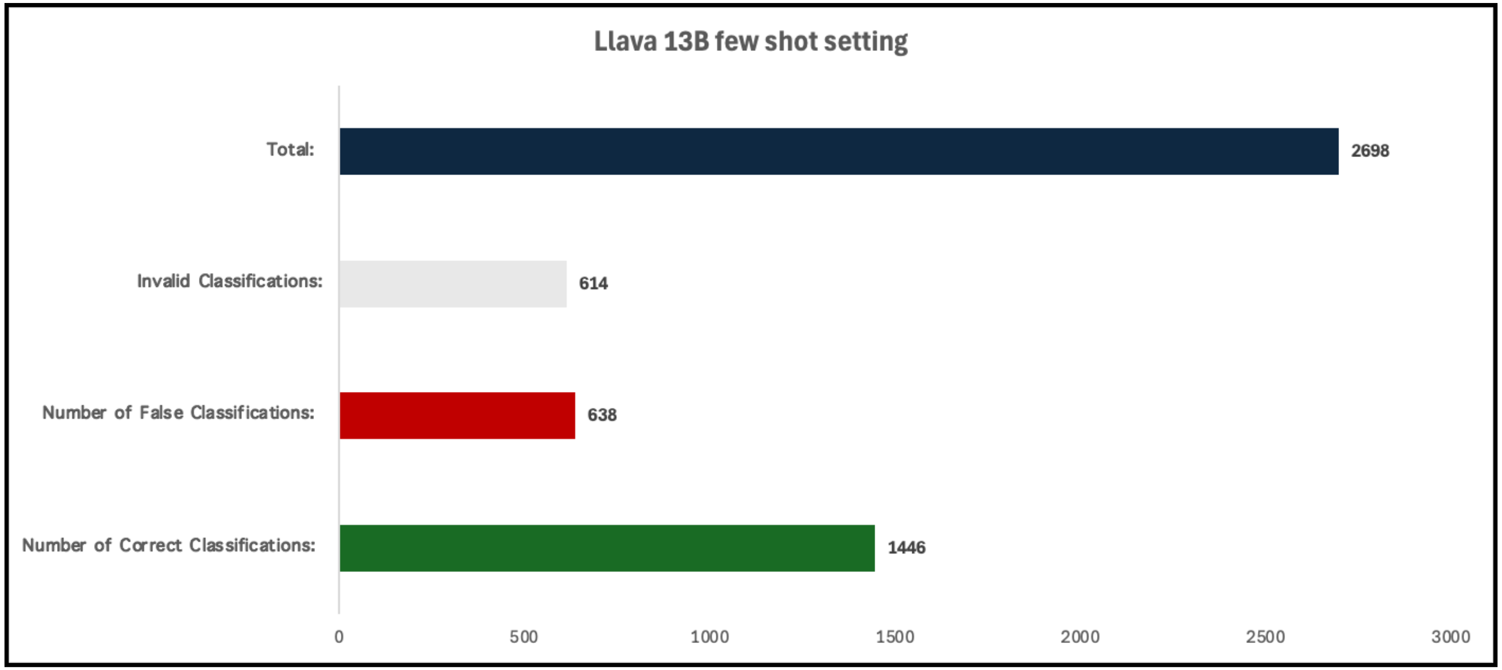
Few Shot Prompting - Bakllava

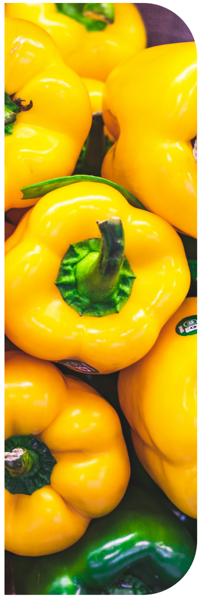


Zero Shot Prompting - Llava 13B



Few Shot Prompting - Llava 13b





Models

Support Vector Machine

Definition

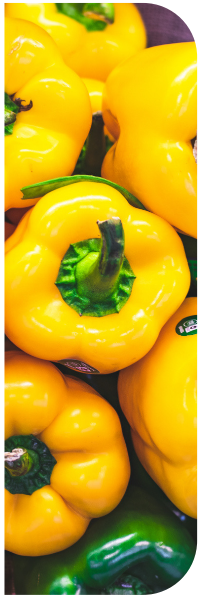
A machine learning algorithm that classifies data by identifying a hyperplane in an N-dimensional space, which separates data points into distinct classes while maximizing the margin between the classes.

Advantages

- Allows user to manually feature engineer
- Robust to overfitting with proper parameter tuning
- Higher interpretability than CNNs, for example
- Suitable for high-dimensional data through kernel tricks, ideal for images with many features
- Doesn’t require as much data

Disadvantages

- Hyper parameters require careful attention to properly tune (kernel, for example)
- Manual feature extraction is more time-consuming
- Kernel functions can be computationally expensive
- Sensitive to noise
- Not suitable for complex or large datasets



Smaller dataset (13,000)

Accuracy: 94.81%

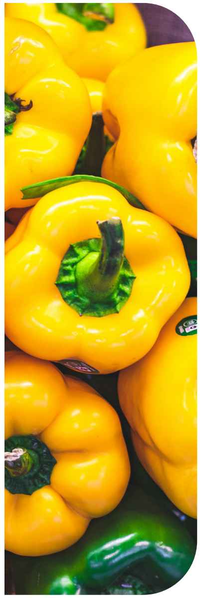
Classification Report:

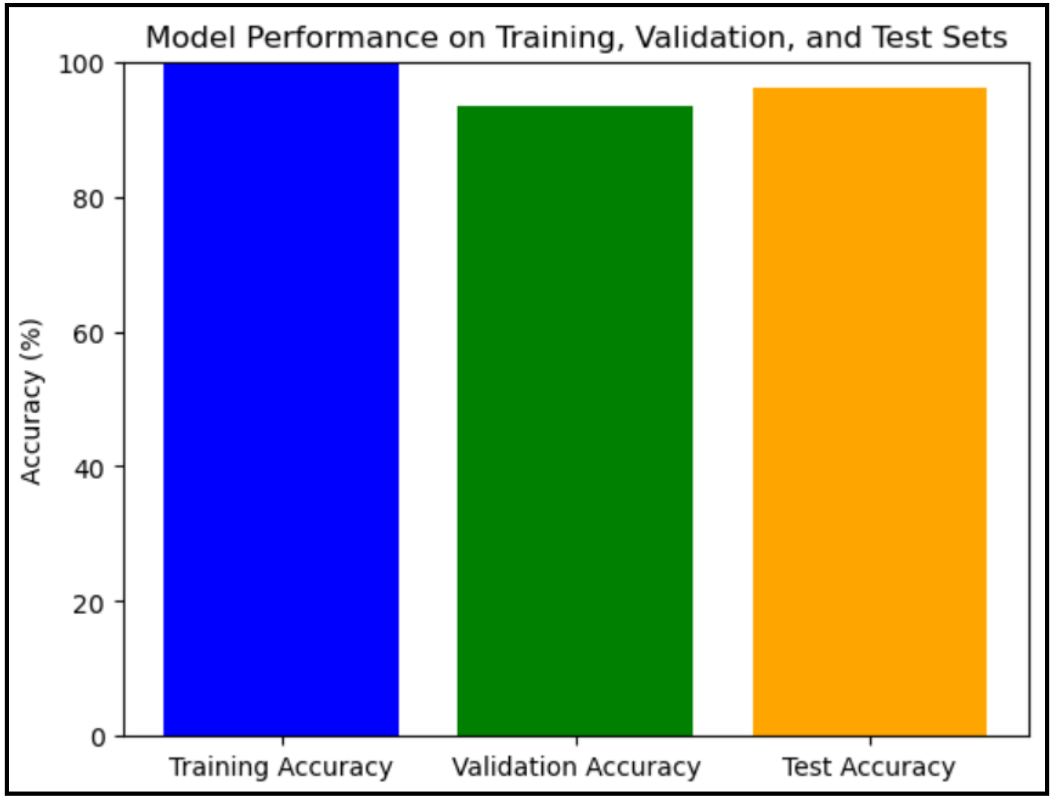
	precision	recall	f1-score	support
0	0.94	0.94	0.94	395
1	0.96	0.98	0.97	381
2	0.95	0.93	0.94	388
3	0.92	0.95	0.93	601
4	0.99	0.97	0.98	530
5	0.93	0.91	0.92	403
accuracy			0.95	2698
macro avg	0.95	0.95	0.95	2698
weighted avg	0.95	0.95	0.95	2698

Larger dataset (30,000)

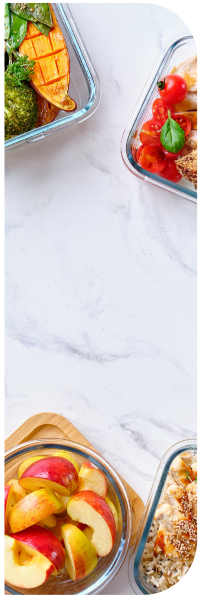
Classification Report (Test Set):

	precision	recall	f1-score	support
freshapples	0.97	0.97	0.97	791
freshbanana	0.99	0.99	0.99	892
freshcucumber	0.96	0.96	0.96	279
freshokra	0.96	0.97	0.96	370
freshoranges	0.95	0.91	0.93	388
freshpotato	0.95	0.93	0.94	270
freshtomato	0.99	1.00	0.99	255
rottenapples	0.93	0.97	0.95	988
rottenbanana	0.99	0.98	0.99	900
rottencucumber	0.96	0.94	0.95	255
rottenokra	0.96	0.97	0.97	224
rottenoranges	0.92	0.86	0.89	403
rottenpotato	0.94	0.96	0.95	370
rottentomato	0.99	1.00	0.99	353
accuracy			0.96	6738
macro avg	0.96	0.96	0.96	6738
weighted avg	0.96	0.96	0.96	6738






```
Training Accuracy: 99.98%
Validation Accuracy: 93.54%
Test Accuracy: 96.26%
```



Models

Contrastive Learning

Definition

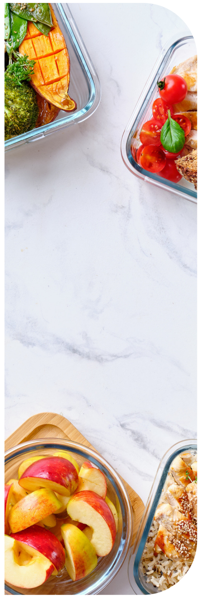
Focuses on extracting meaningful representations from data by contrasting positive and negative pairs of instances. This approach aims to map similar instances close together in a latent space while pushing dissimilar instances apart.

Advantages

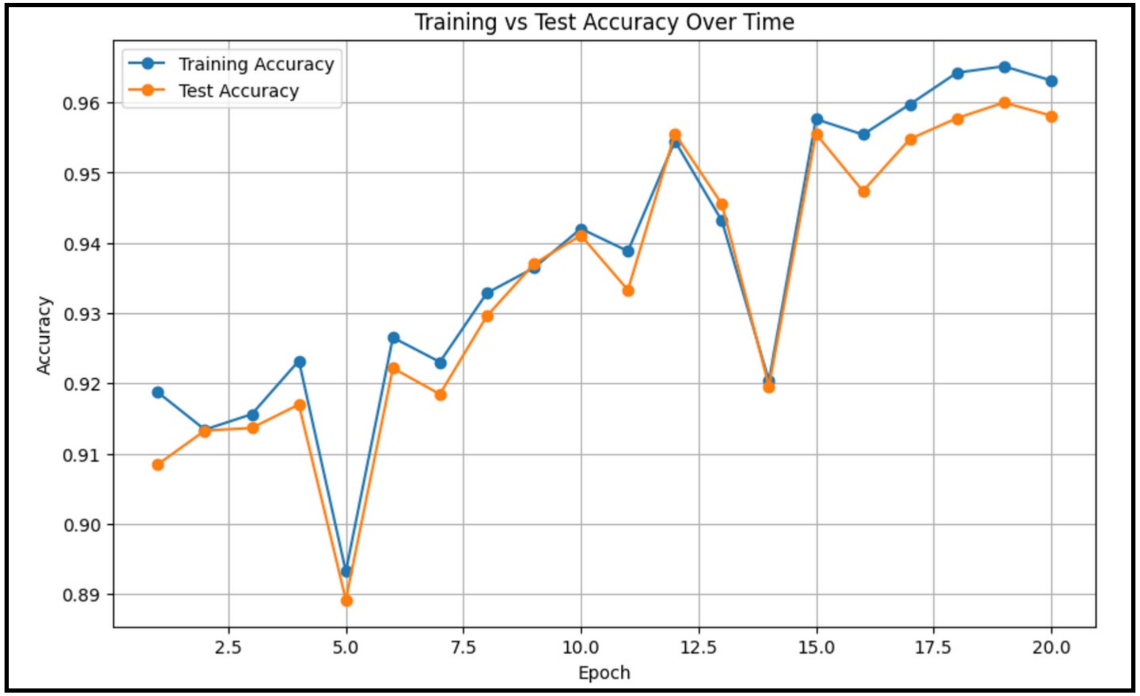
- Ability to learn unlabelled data makes it exceptionally beneficial
- By leveraging similarity and dissimilarity, contrastive learning enables models to capture relevant features and similarities in the data

Disadvantages

- Complex and can require significant computational power
- Could lead to biased representations
- Challenging to tune and choose negative samples



Contrastive Learning



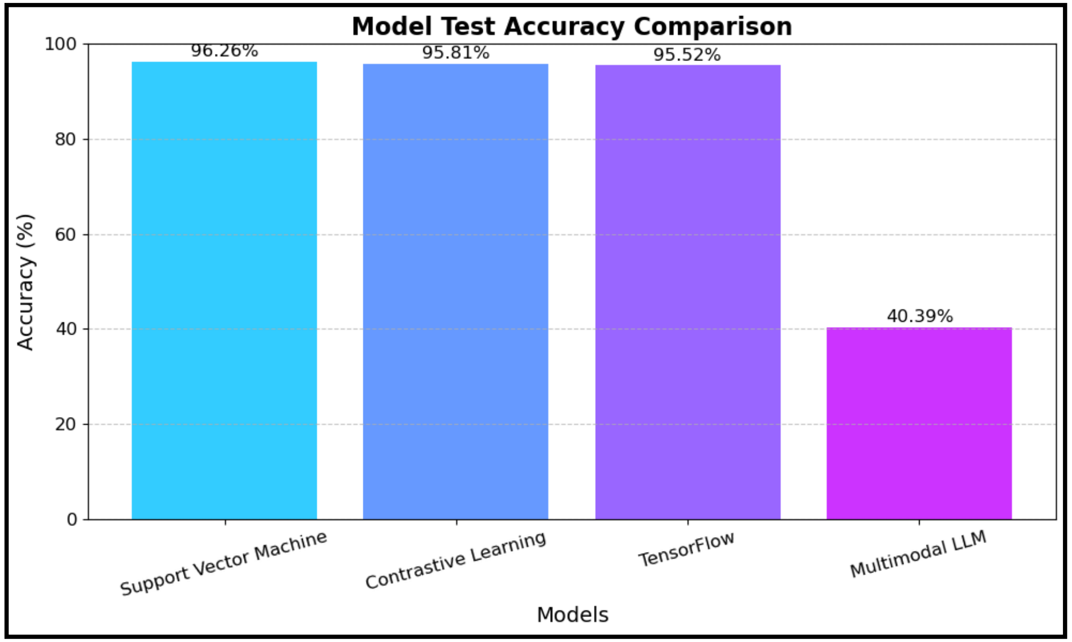
LEARNING_RATE = 0.0001
NUM_EPOCHS = 20
BATCH_SIZE = 32
BEST TEST ACCURACY: 0.9589

Results

Comparing Models

The best model in terms of highest accuracy on the test set is:

Contrastive learning is great for learning meaningful feature representations by comparing similar and dissimilar data, while SVMs work well with those extracted features or when the differences between classes are clear.



Beyond the Models

Applications

Quality Control & Assurance

Automated systems can play a pivotal role in ensuring the quality of fruits and vegetables by efficiently identifying and sorting fresh produce from spoiled or damaged items

Reducing Food Waste

Identified spoiled items can be diverted to composting or sold at a discount before they become inedible, maximizing their value and minimizing waste

Smart Refrigeration Systems

Incorporating freshness detection technology into smart refrigerators can alert users when items are nearing spoilage, encouraging timely consumption

Disease Detection in Crops

Early detection of spoilage can also serve as an indicator of potential disease outbreaks in crops, allowing for swift intervention to prevent further contamination

Wrap-Up

Concluding Thoughts

Novel Aspect of Our Project

Multimodal LLMs
Contrastive Learning

Best Models

Support Vector Machine
(Suitable for binary classification, robust to overfitting)

Contrastive Learning
(Focuses on subtle features to classify images)

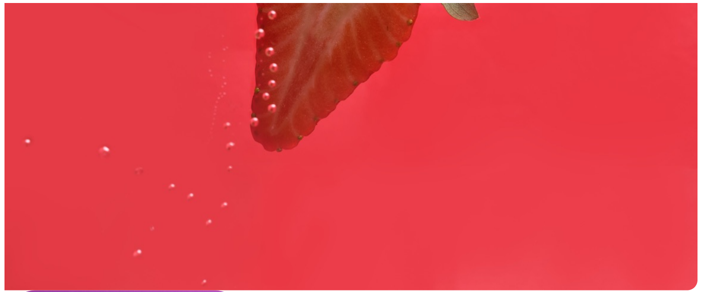
What We Learned

Diversity of models
Some models handle images and binary
data better than others

Q&A

Questions?

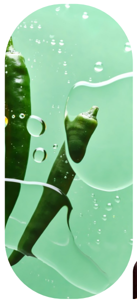
Thank you for your attention! We're excited to answer any questions you may have.

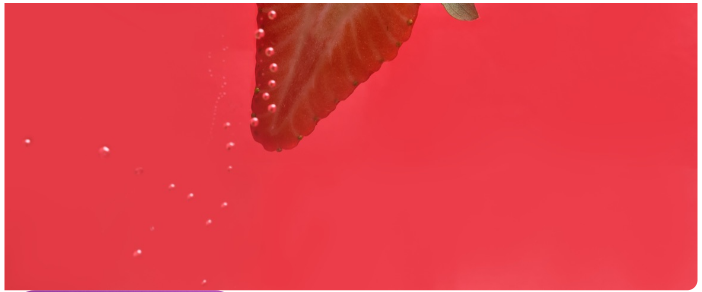


Closing

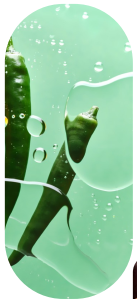
Thank You!



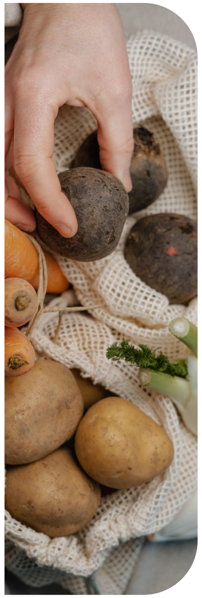




Appendix


This just shows that we tried this model but were finding very strange results, so we focused our attention to the superior models



Models

ResNet-50 CNN

Definition

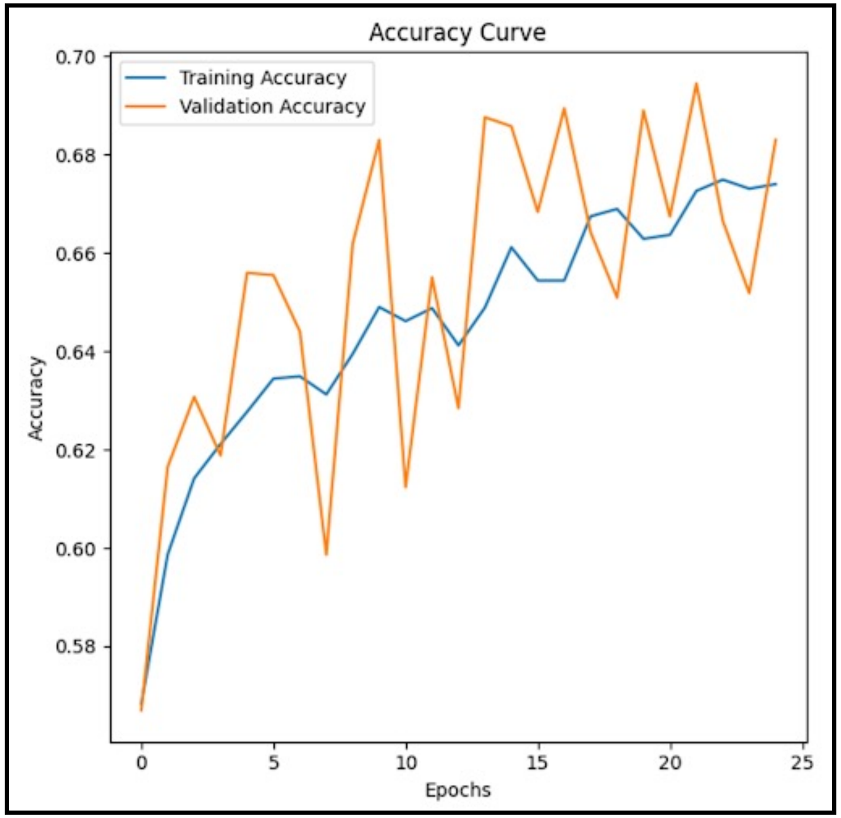
NN architecture specifically designed for image classification tasks. By using residual connections, it can overcome the vanishing gradient problem.

Advantages

- Residual connections
- Efficient computation
- Faster convergence

Disadvantages

- Large model size
- Overfitting on small datasets
- Not the most efficient architecture



Test Accuracy: 73.20%
Learning rate: 0.001
Epochs: 25

This just shows that we tried this model but were finding very strange results, so we focused our attention to the superior models

Models

EfficientNet B3 CNN

Definition

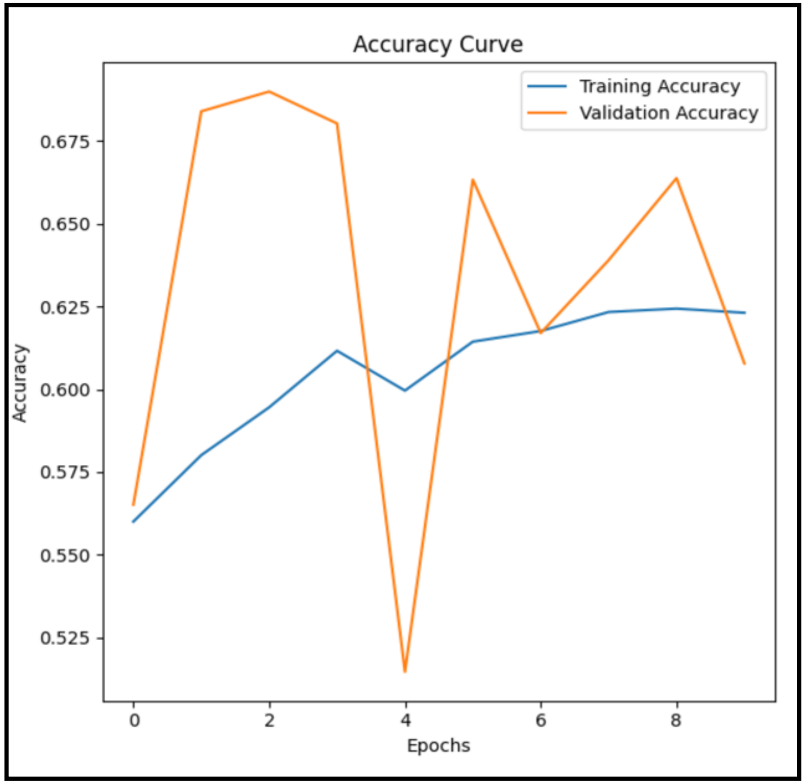
NN architecture that balances computational efficiency and model performance through compound scaling method

Advantages

- Better accuracy than older models
- Efficient use of parameters
- Versatility

Disadvantages

- Slower training
- Dependency on pre-trained models
- Computational complexity for larger variants



Test Accuracy: 67.90%
Learning rate: 0.001
Epochs: 25