

PERFORMANCE BENCHMARKING OF YOLO ARCHITECTURES

For Real-Time Object Detection

Beyza GÜLER | YOLOv11 Specialist

Ramazan YILDIZ | AI Project Lead

Abdelrahman MOHAMED | Web Developer

Presentation Agenda

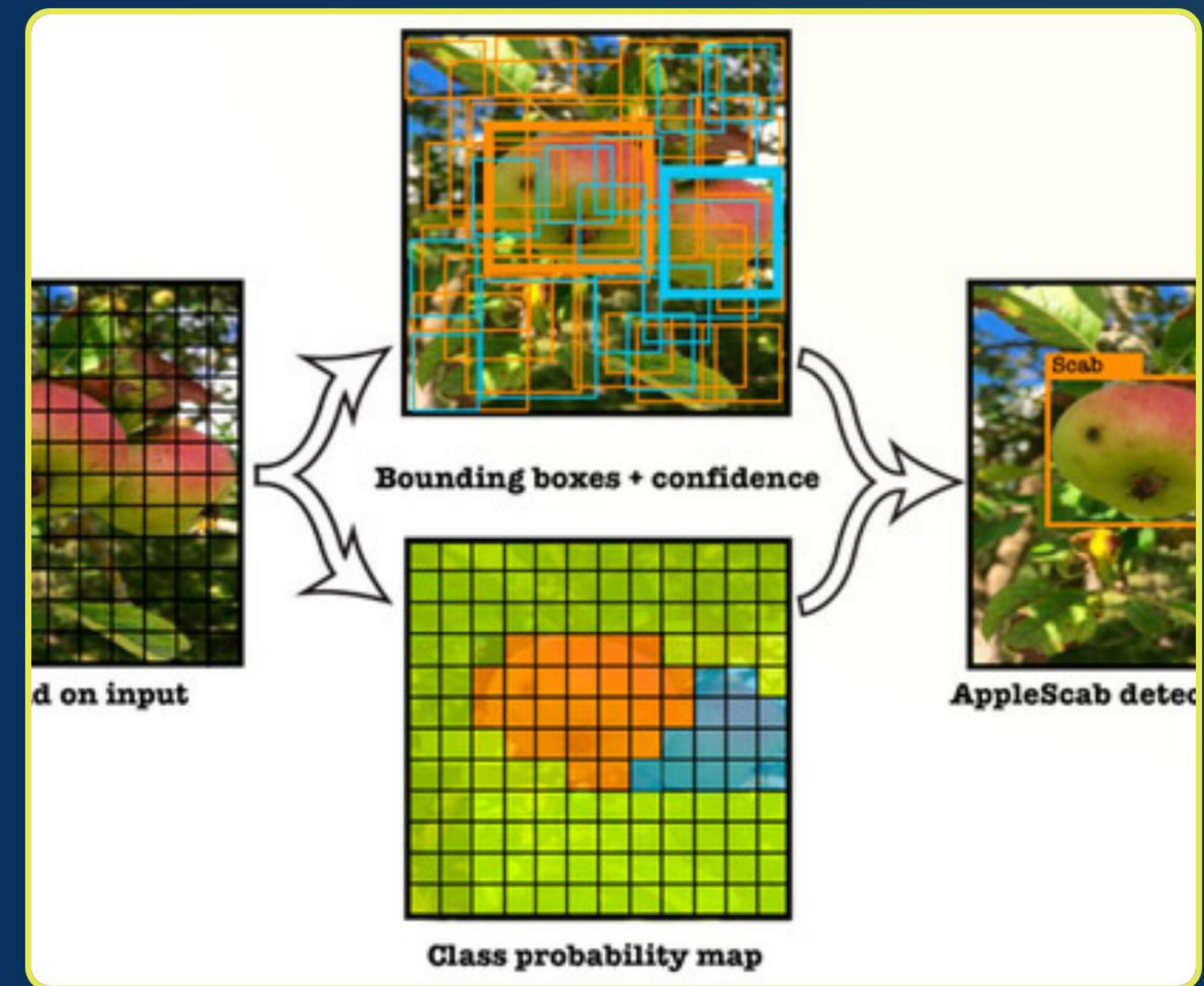
- ✓ Project Abstract & Goals
- ✓ Methodology & System Pipeline
- ✓ Dataset Overview (Roboflow)
- ✓ YOLO Model Architecture
- ✓ Training Analysis & Results
- ✓ Benchmark Comparison (v5 vs v8 vs v11)
- ✓ Deployment & Web Interface
- ✓ Conclusion & Future Work

Project Abstract

Goal: Real-time fruit detection using YOLO models via a web interface.

Core Tech: Pre-trained YOLO + Transfer Learning + Streamlit Deployment.

Outcome: A working prototype achieving stable ~30 FPS.



Meet the Team



Beyza GÜLER

YOLOv11 Specialist

Analysis & Reporting



Ramazan YILDIZ

AI Project Lead

Research & Planning



**Abdelrahman
MOHAMED**

Web Developer

YOLOv5 & Frontend

The Problem

The Trade-off

High accuracy models are often too slow (low FPS) for real-time applications.



Speed vs. Accuracy

Deployment Gap

Complex AI models are difficult to access without a user-friendly web interface.

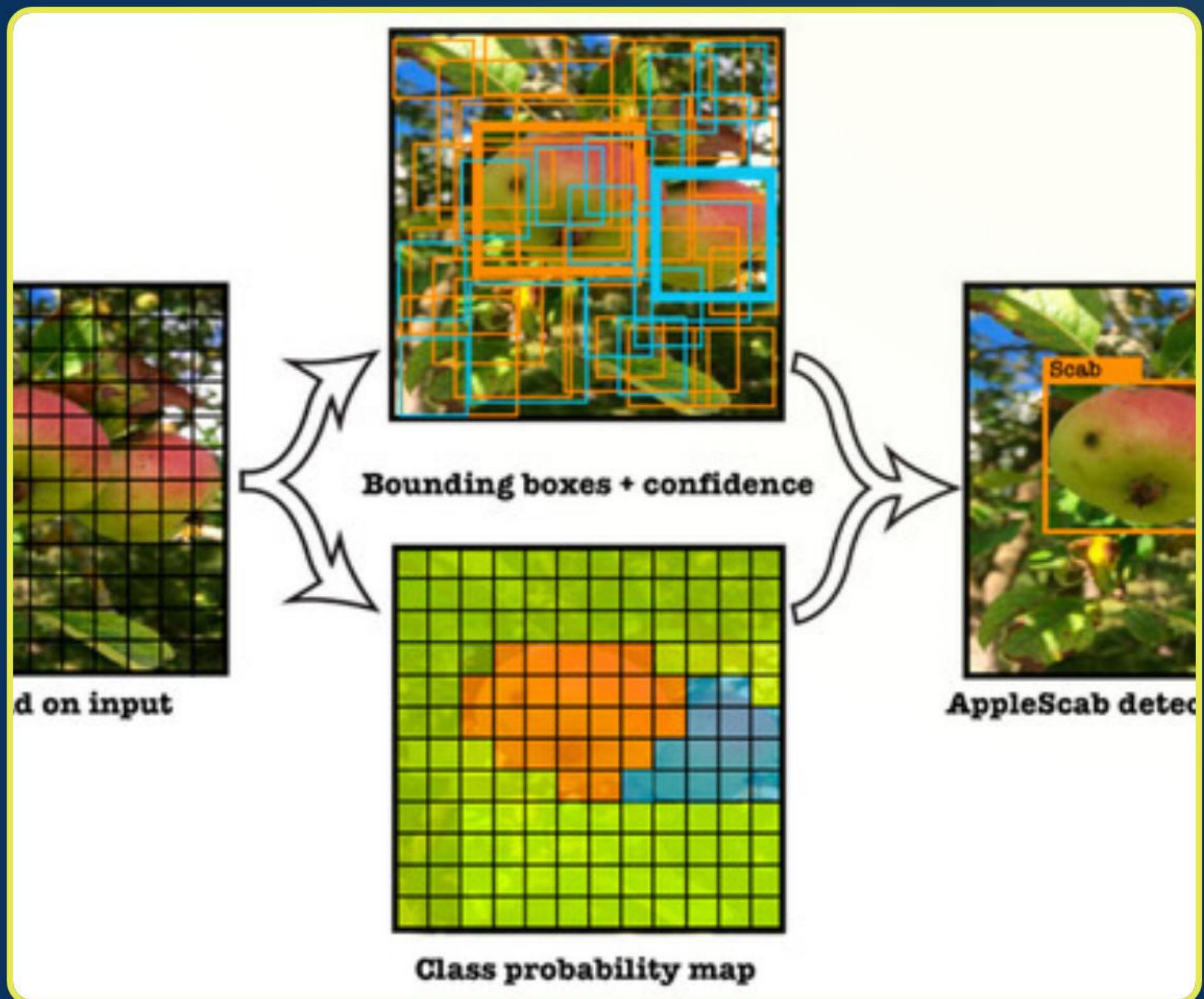
Project Goals

- ✔ **Achieve Real-Time Performance:** Target ~30 FPS for smooth video.
- ✔ **Develop Web Interface:** Accessible Streamlit app for end-users.
- ✔ **Benchmark Architectures:** Compare YOLOv5, YOLOv8, and YOLOv11.
- ✔ **Scalability:** Create a modular system ready for future expansion.

What is Object Detection?

Unlike simple image classification (which labels a whole image), **Object Detection**:

- ✓ Locates objects within the frame.
- ✓ Draws bounding boxes around them.
- ✓ Classifies each box with a confidence score.



Why Choose YOLO?



Single-Stage

Detects objects in a single forward pass, making it incredibly fast.



Efficient

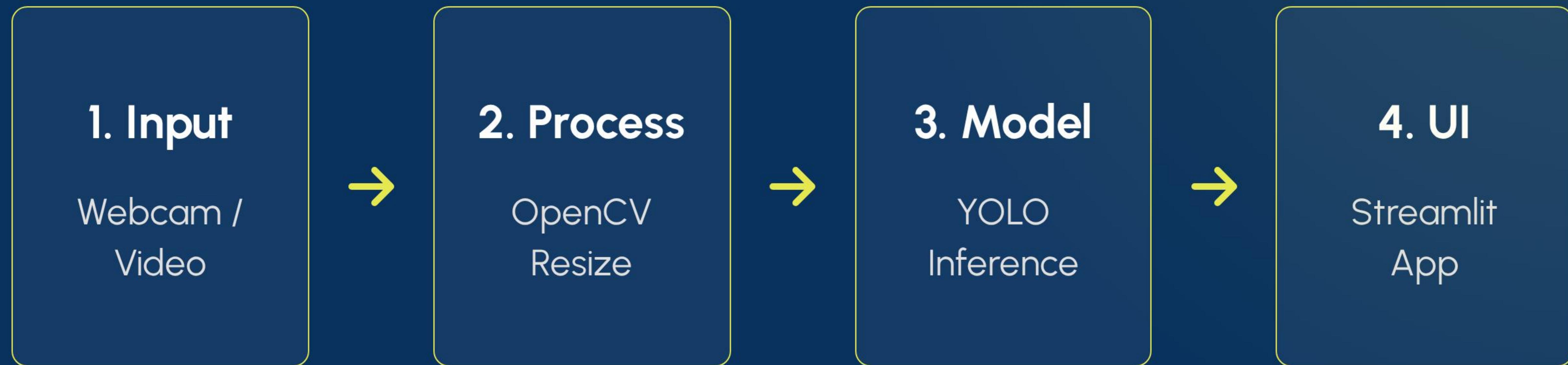
Balances speed and accuracy better than two-stage detectors like R-CNN.



Community

Strong support, pre-trained weights, and active development.

System Pipeline



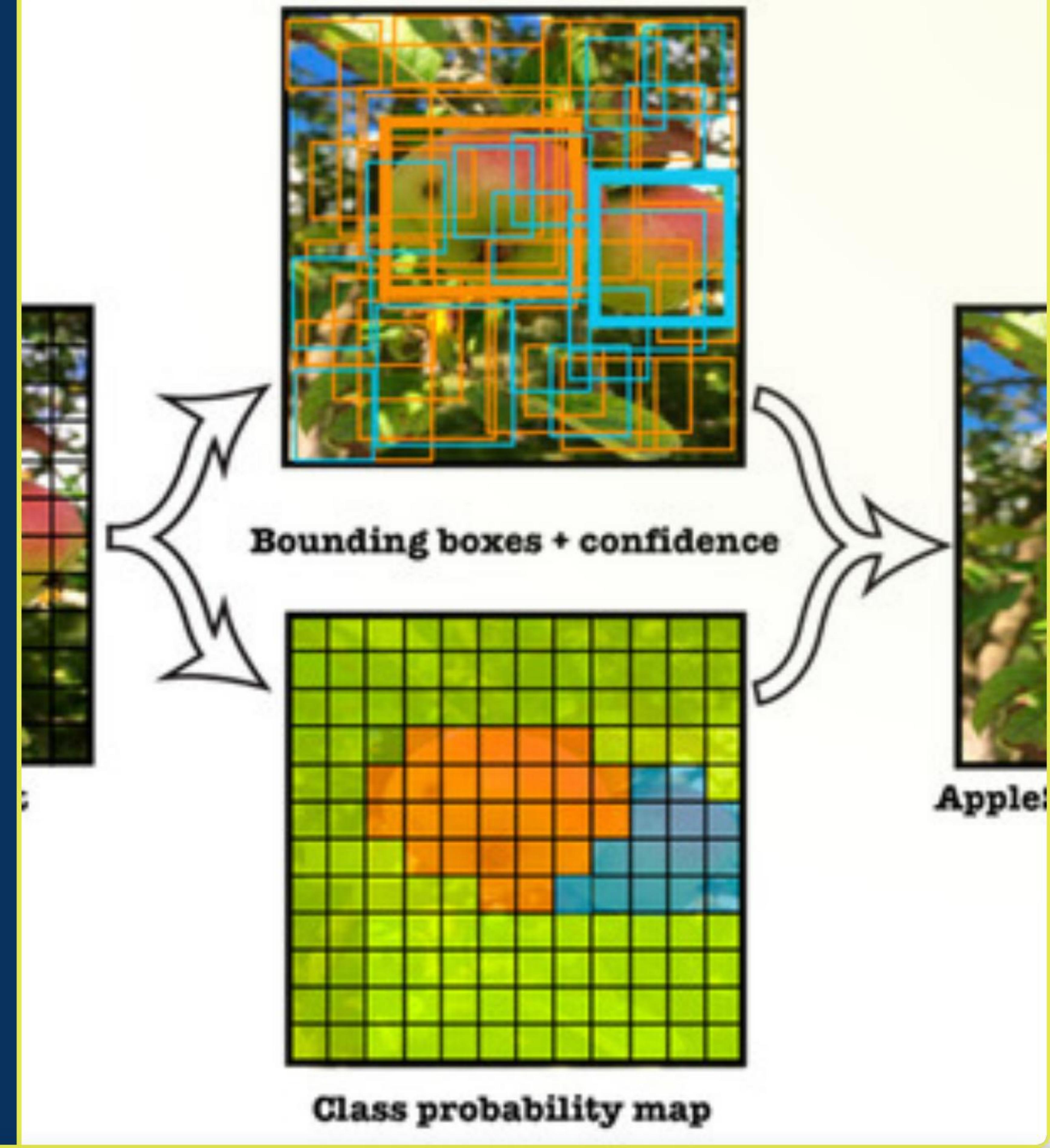
The Dataset

Source: Roboflow Fruit Detection (v1)

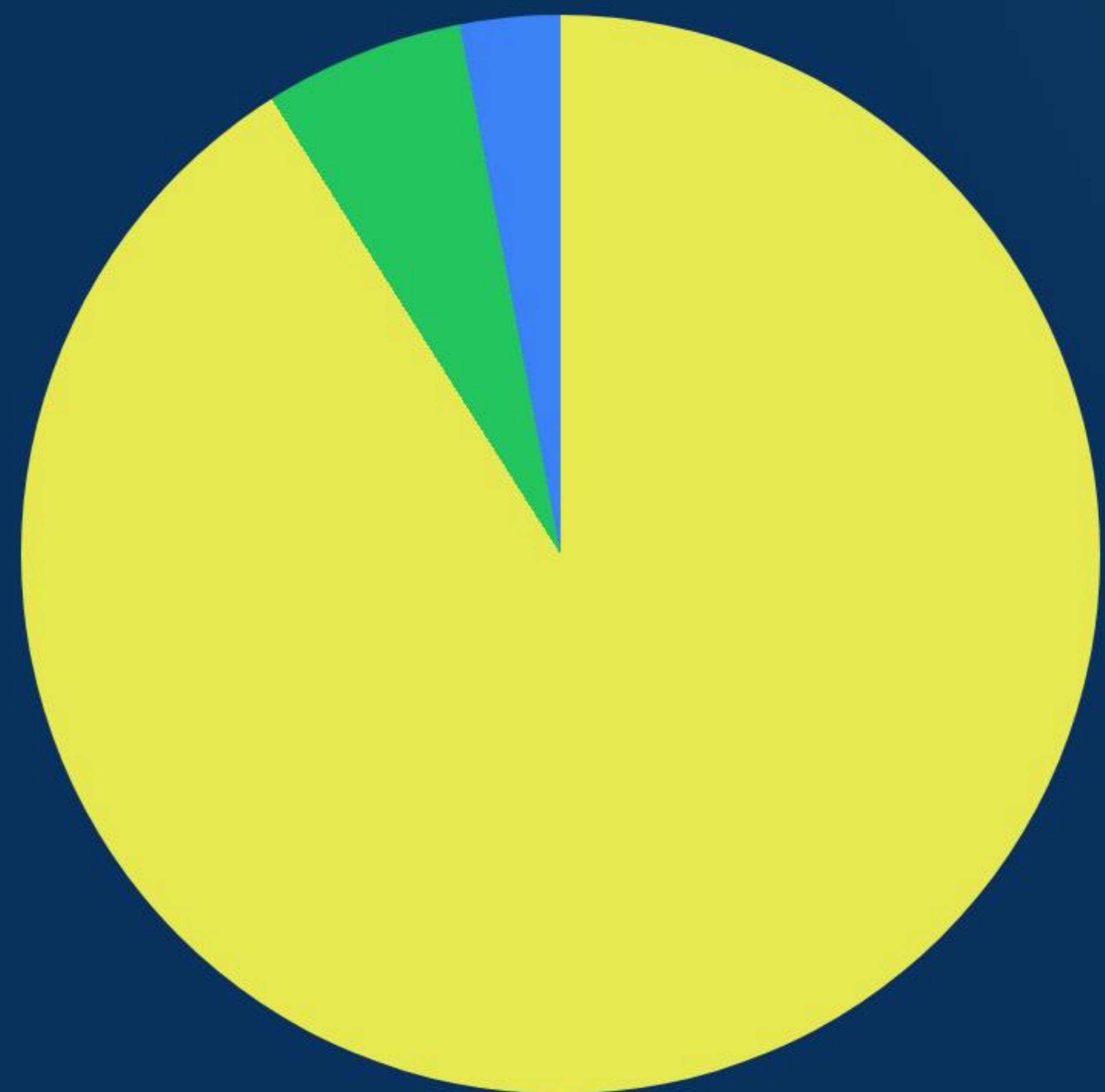
Total Images: 2,974

Classes: 9 distinct fruit types

Pre-labeled and normalized for immediate YOLO training.



Dataset Split



Total: 2,974 Images



Architecture

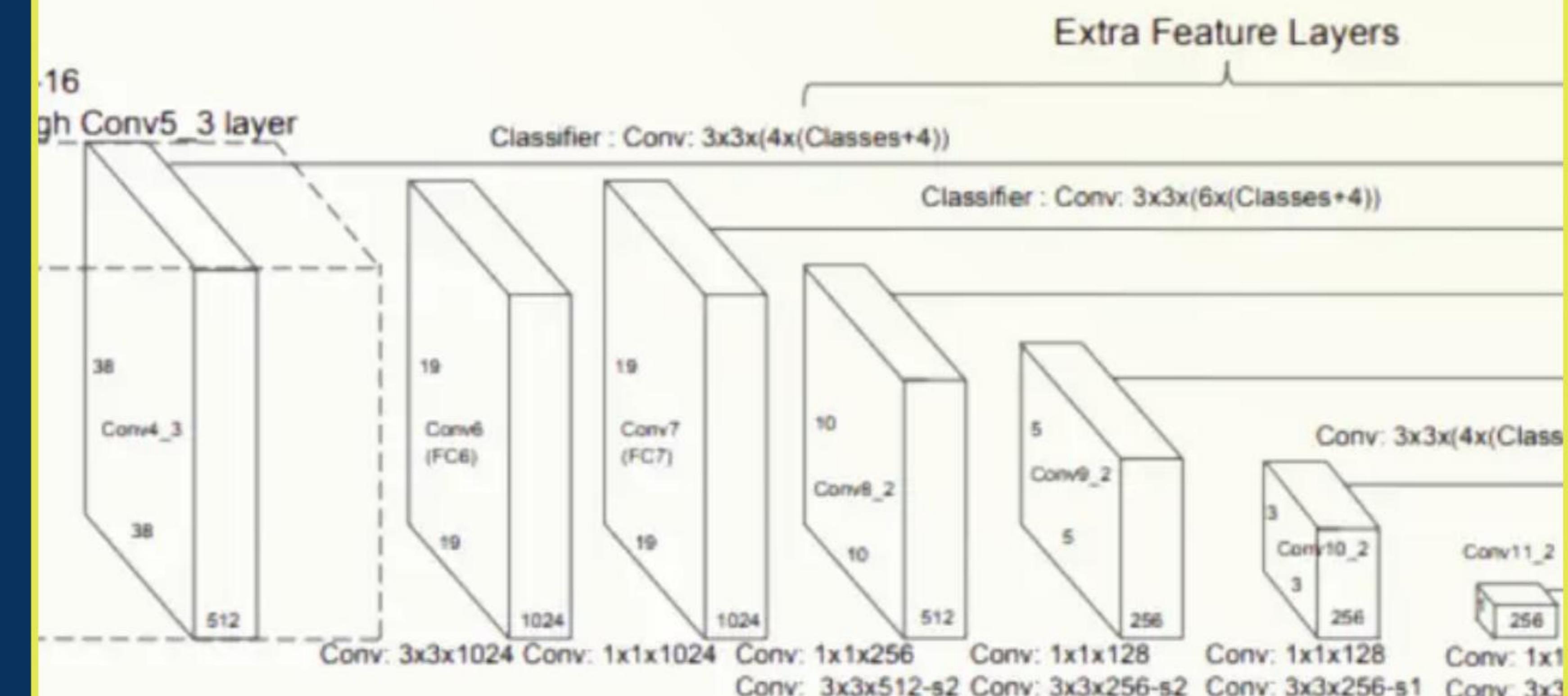
Backbone: Extracts features from the input image.

Neck: Aggregates features across different scales.

Head: Makes final predictions (Bounding Box + Class).

Single-stage design ensures real-time performance.

Structure of SSD



Training Setup

Environment: Google Colab (GPU)

Framework: Ultralytics YOLO (PyTorch)

Hyperparameters:

- ✓ Epochs: 50
- ✓ Batch Size: 16
- ✓ Image Size: 640px

```
# data.yaml configuration
train: ../train/images
val: ../valid/images
nc: 9
names: ['apple', 'banana', ...]
```

Training Analysis

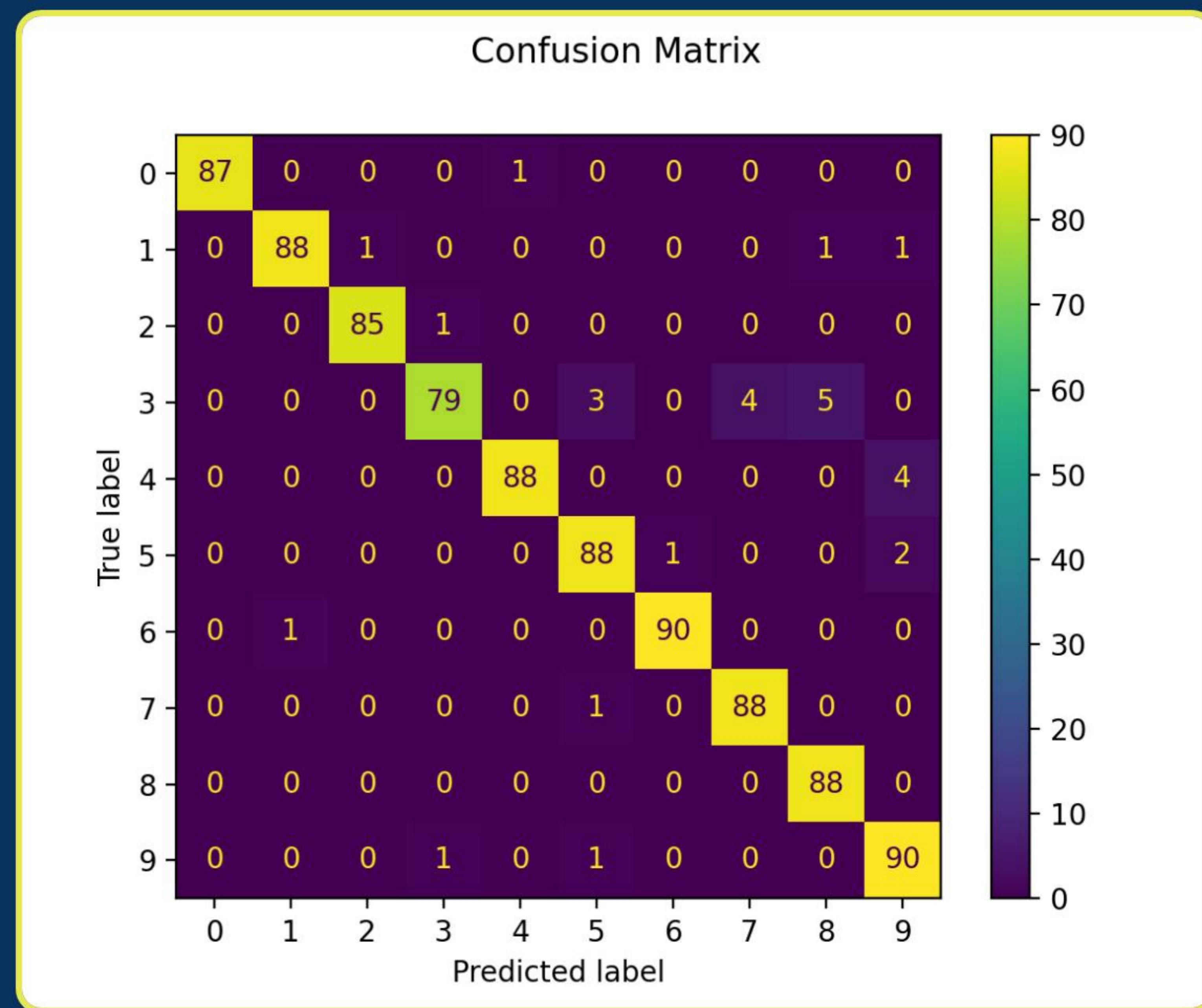
Key Observations

Loss Curves: Showed consistent decrease, indicating successful learning.

mAP Scores: Steadily increased over 50 epochs.

Stability: No significant signs of overfitting were observed.

Confusion Matrices



Strong Performance

The matrix shows a strong diagonal concentration.

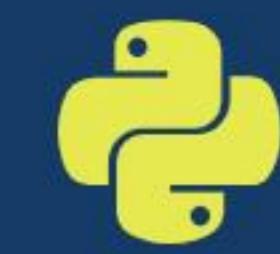
This confirms that the models correctly classify fruit types with minimal confusion between classes.

Benchmark Results

Model	Accuracy Strength	Speed Strength
YOLOv11m	Highest Accuracy	Moderate Speed
YOLOv5m	Good Accuracy	Fastest Speed
YOLOv8m	Balanced	Balanced

Target Met: All models achieved ~30 FPS Real-Time Performance.

Deployment Stack



Python

Core Logic



OpenCV

Preprocessing



Streamlit

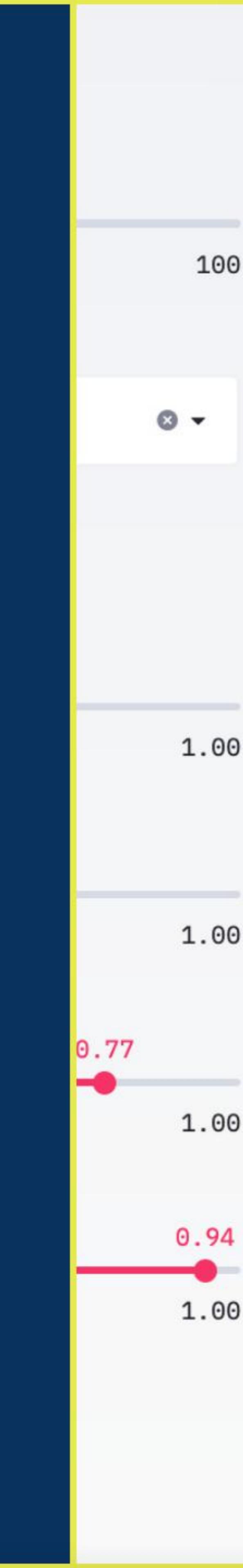
Web Interface

Web Interface

User-friendly browser application.

Features:

- ✓ Live Webcam Feed.
- ✓ Real-time Bounding Boxes.
- ✓ Live Confidence Scores.

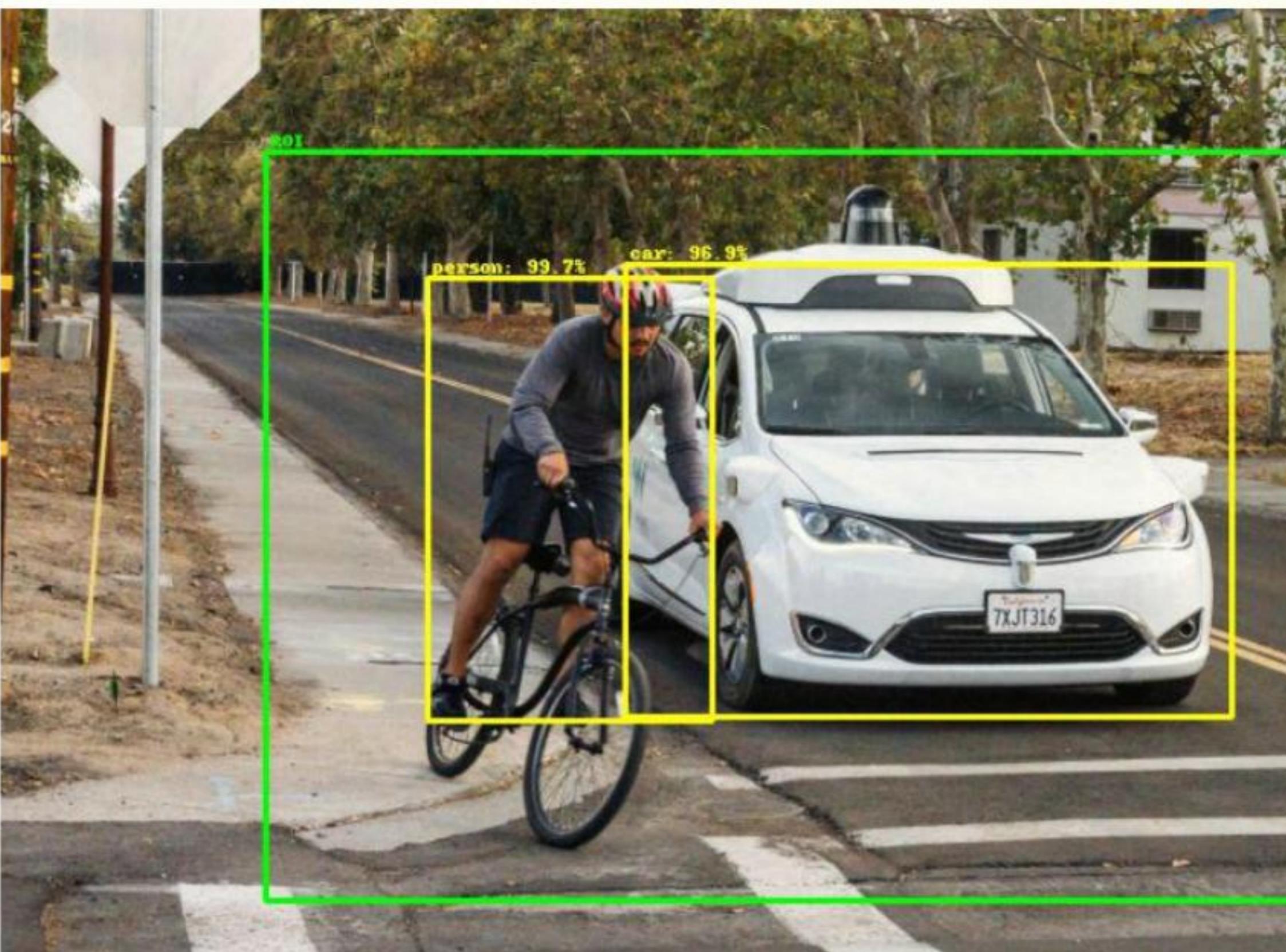


Deepstack Object detection

Upload an image

street.jpg

browse files



All discovered objects

```
{'stop sign', 'car', 'person', 'bicycle'}
```

Filtered object count

car : 1

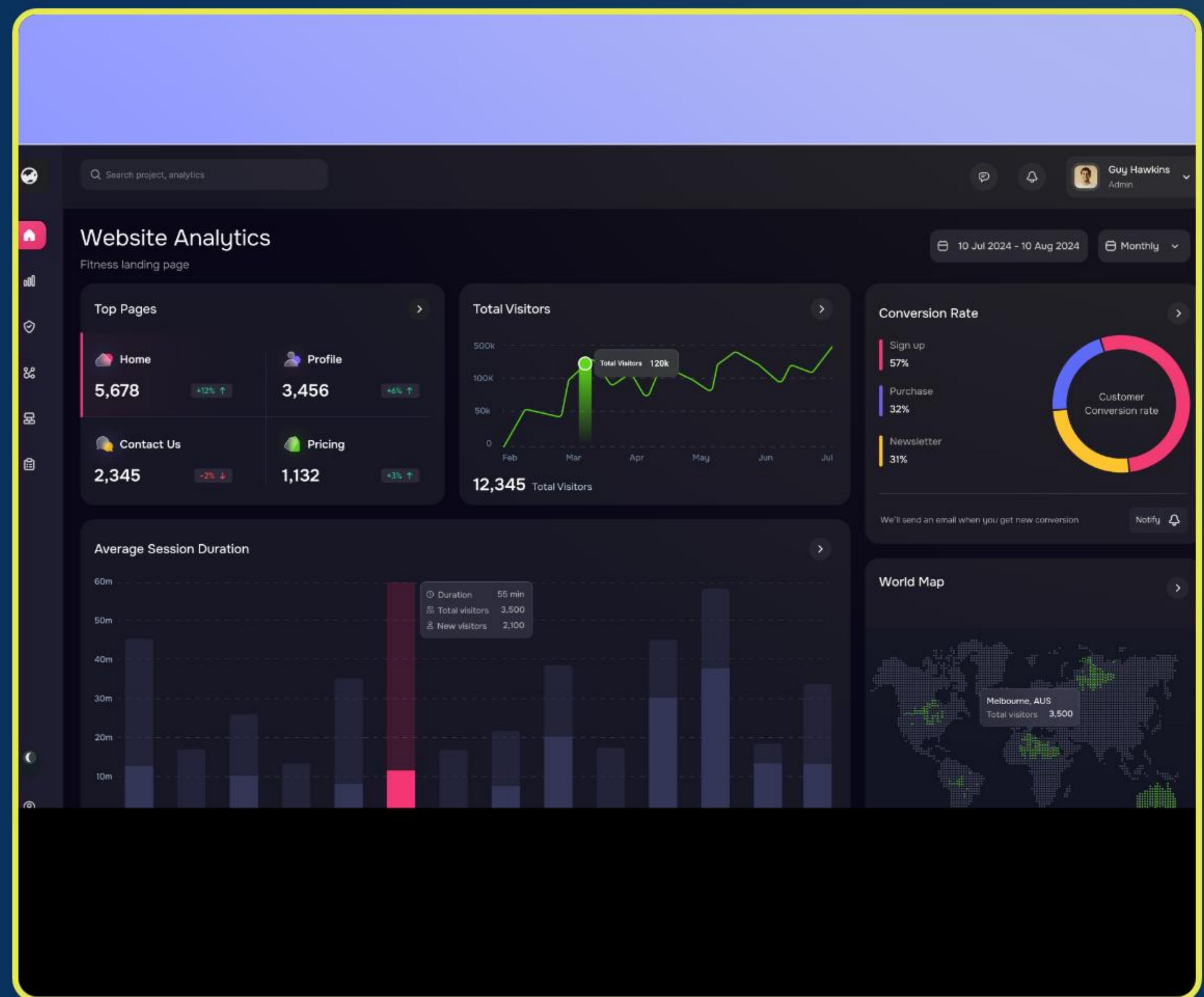
person : 1

Analytics Dashboard

Data Insights

The system tracks key metrics in real-time:

- Total Detections Count
- Average Confidence Score
- Per-Class Distribution



Future Roadmap



Dataset

Expand to 10k+ images
for robustness.



Mobile App

Deploy to Android/iOS
for portability.



3D Tech

Implement 3D
localization & depth.



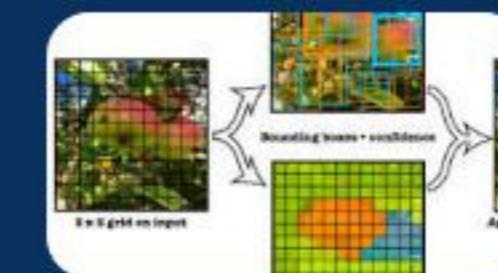
MISSION ACCOMPLISHED

Stable, Real-Time Detection at 30 FPS

Thank You!

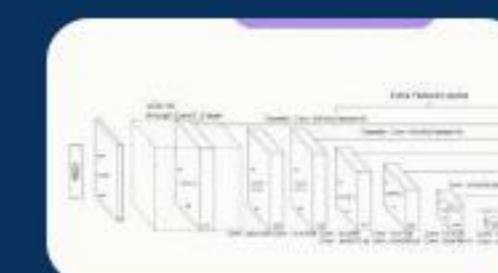
Beyza GÜLER | Ramazan YILDIZ | Abdelrahman MOHAMED

Image Sources



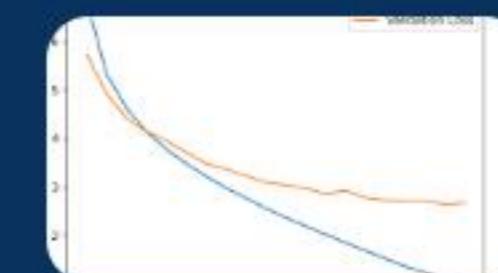
https://www.mdpi.com/blsf/blsf-30-00006/article_deploy/html/images/blsf-30-00006-g001-550.jpg

Source: www.mdpi.com



<https://a.storyblok.com/f/139616/1200x800/89f31e46c8/structure-of-ssd.webp>

Source: encord.com



https://machinelearningmastery.com/wp-content/uploads/2022/10/training_validation_loss_1.png

Source: machinelearningmastery.com



https://miro.medium.com/v2/resize:fit:1400/1*tgSTfVKWof4VDiAi_HjTQg.png

Source: organized-curiosity.medium.com



<https://global.discourse-cdn.com/streamlit/original/2X/4/4f84daa35cdbcff5d35f0a8597faf76ad8639d14.jpeg>

Source: discuss.streamlit.io



<https://cdn.dribbble.com/userupload/43387713/file/original-6af3358e1ae38eab78e7853ab5e95798.png>

Source: dribbble.com