

Computer Vision for Intelligent Pest Detection

AutoEncoder + YOLOv8

Sathwick

Problem

Pests significantly impact crop health, yield, and quality.

Challenges in Pest Detection:

1. Noise
2. Subtle Differences

Goal

An approach that can accurately identifies pests while overcoming challenges of the noise and subtle differences.

Leverage Computer Vision based pest detection techniques.

Dataset

IP102 YOLO:

19,000 Images

- Collected from internet
- 102 kinds of pests
- Bounding Boxes

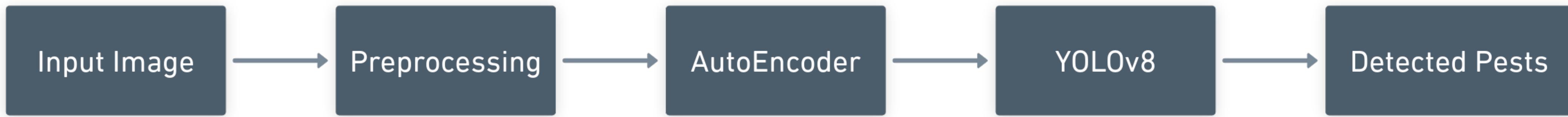
Subset used for developing the Proposed approach:

800 Images

- Randomly chosen images from the existing data
- 10 classes
- 80 images per class

10 Classes - "asiatic rice borer", "mole cricket", "aphids", "beet army worm", "flax budworm", "blister beetle", "Lycorma delicatula", "Miridae", "Prodenia litura", "Cicadellidae"

Approach



Why AutoEncoder?

Two Components

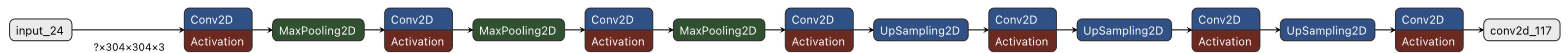
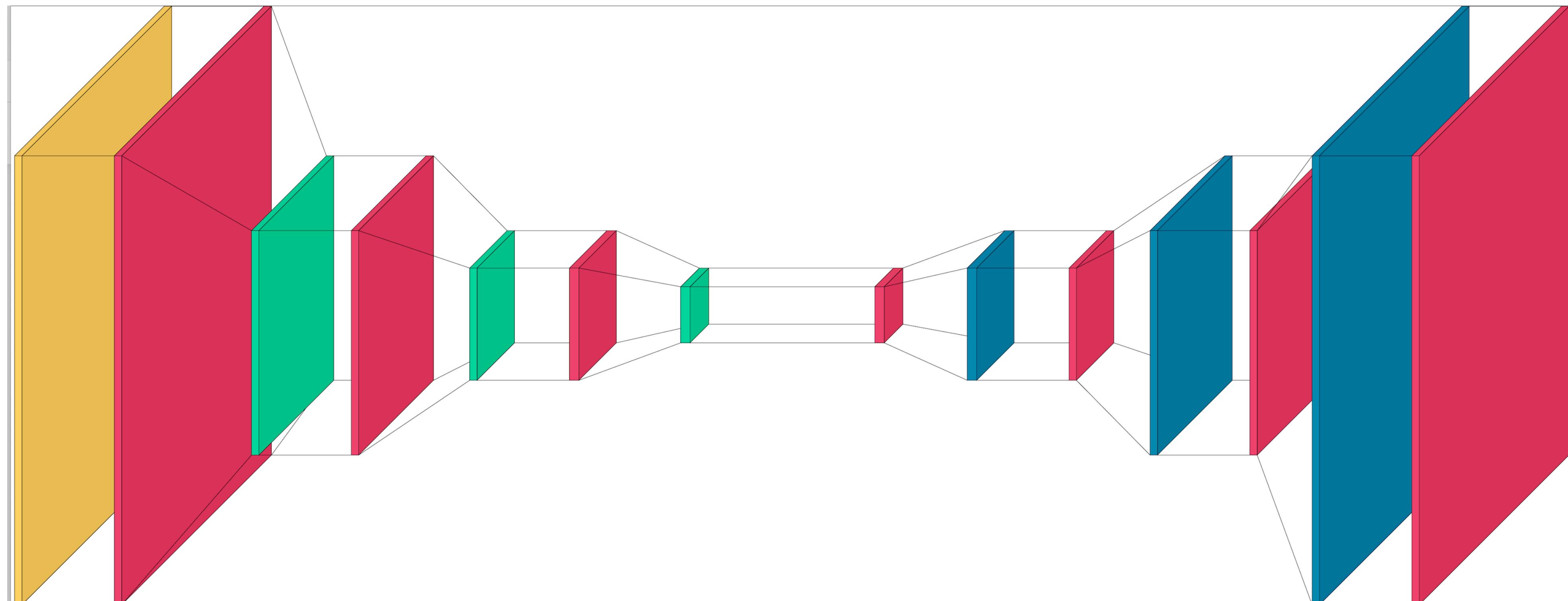
Encoder - Identifies patterns and encodes it in a much smaller dimension

Decoder - Reconstruct the model from the latent representation

Forces AE to learn the important patterns in the image.

Alleviates the problems of Noise and Subtle differences in the images

AutoEncoder



Preprocessing & Learning

Training

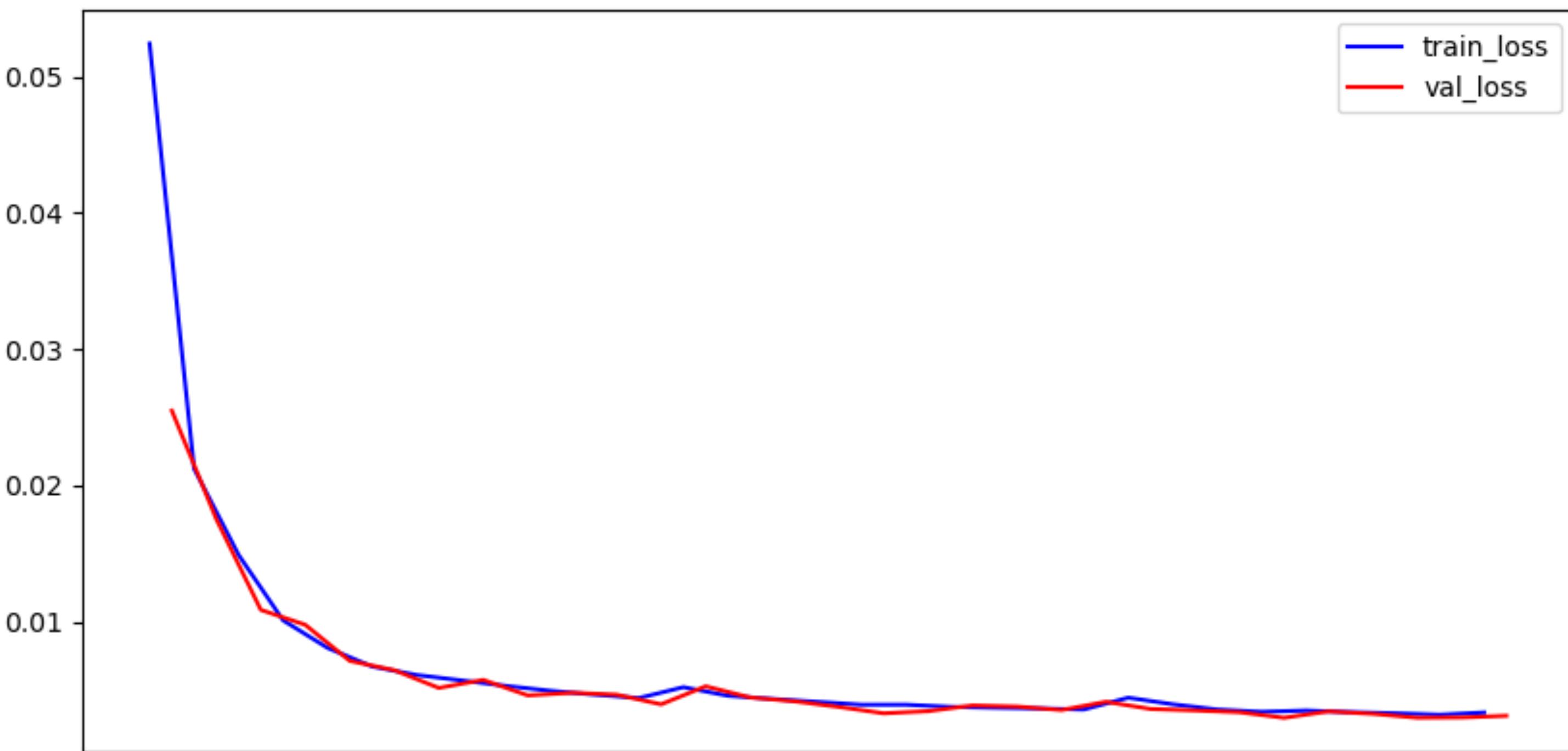
- Input image dimensions selection for uniform shapes
- Add random noise to the image

Learning

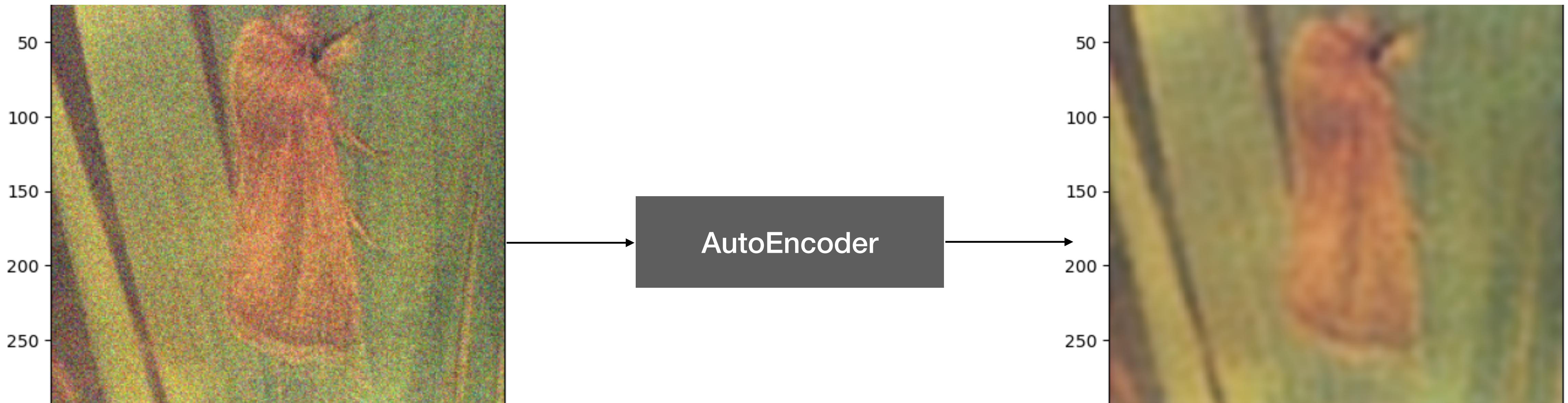
- Keras Sequence and Data Generators

AutoEncoder

MSE on Test: 0.0037



AutoEncoder



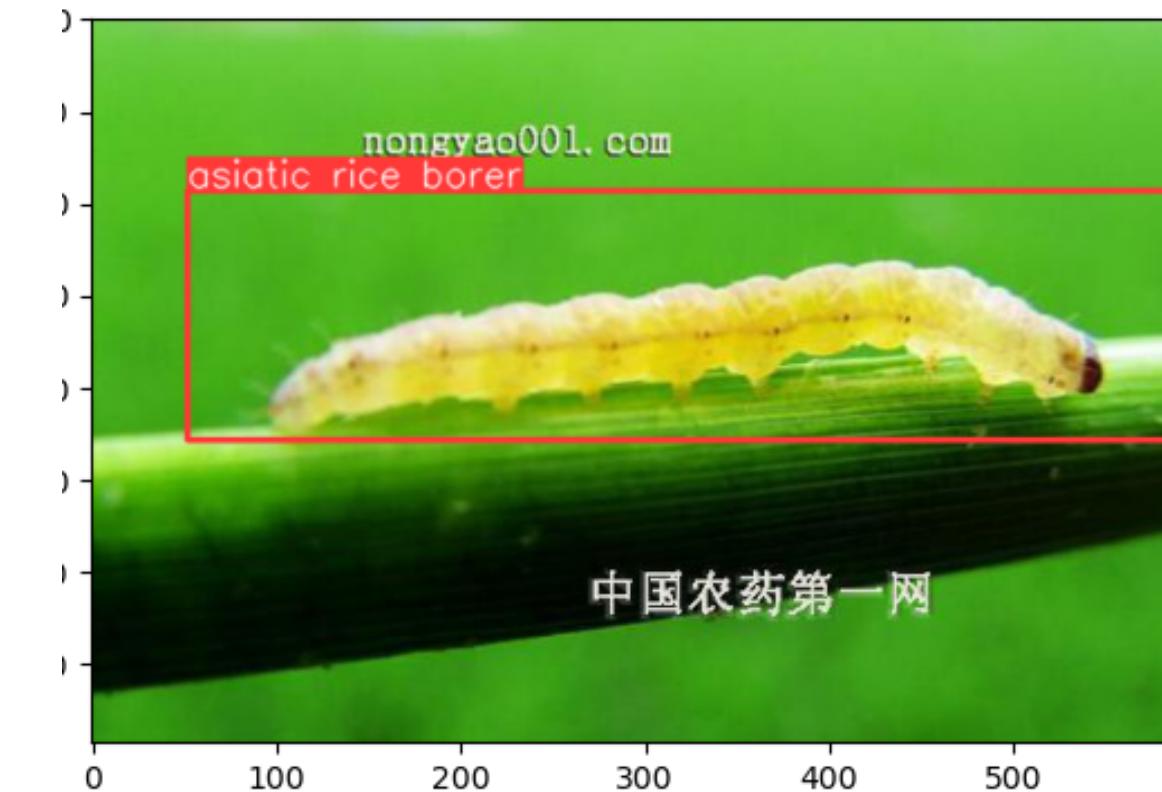
Why YOLOv8?

- Single Pass Inference
- Fast & accurate predictions - suitable for real time predictions
- Adapts to the dynamic scale of lesions
- Precise object localization capabilities

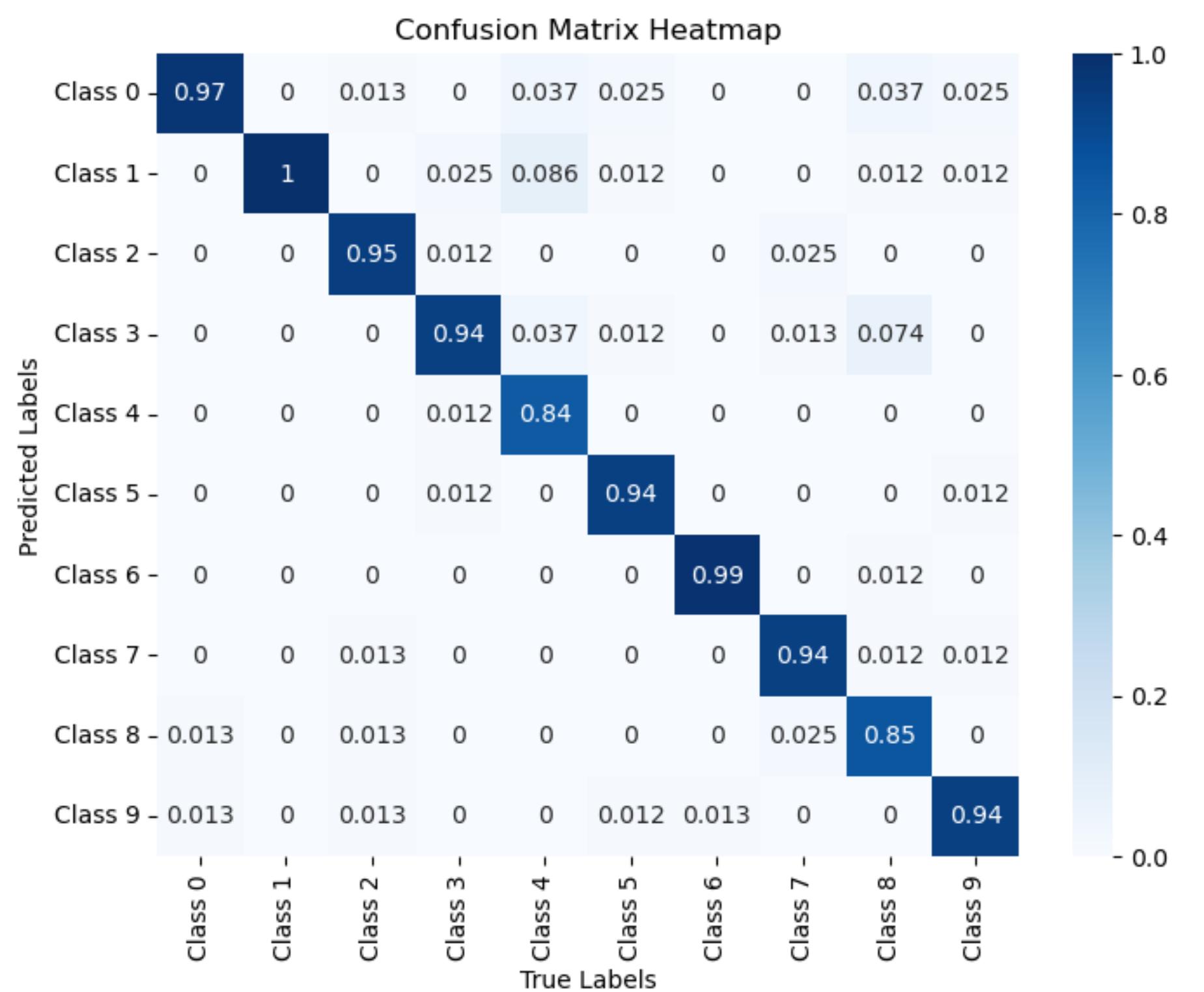
YOLOv8

- Backbone - A modified version of CSPDarknet53 for extracting essential features from the input image
- Convolution Blocks - Basic Convolution, C2F, Bottleneck, SPF
- Head - Follows the backbone with multiple conv layers and FC layers and responsible for bounding boxes and class probability predictions

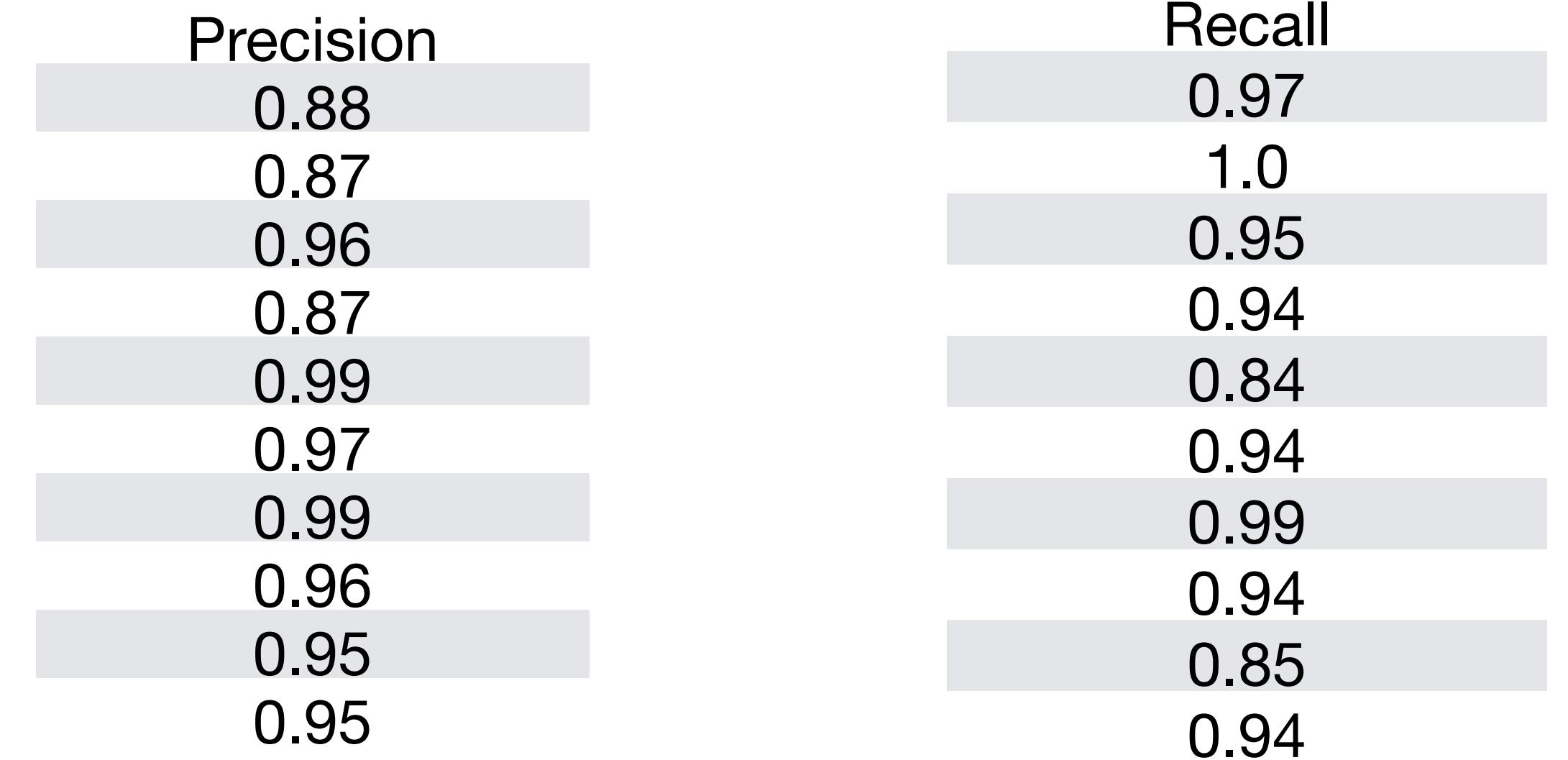
YOLOv8



YOLOv8 Metrics



Accuracy: 93.87%



Further Steps

- Train on the complete dataset to further enhance the accuracy and prediction capabilities.
- Can be easily deployed on to the embedded devices and can be utilized for predictions in real time.
- Possible to further extend to tasks including image segmentation, in video classification

Demo

Demo the functionality and features of the model developed