VineNet: Grasp of the Grapes

(P09-B)

Utilizing semantic segmentation to revolutionize precision viticulture.

The Challenge

Precise Grape Identification

To build a deep learning model capable of accurately segmenting grape bunches (semantic segmentation) in high-resolution images of vineyards.

Why Semantic Segmentation?

Pixel-Level Accuracy

Pinpoint the exact location and boundaries of each grape.

Unlock Valuable Insights

Enable data-driven decision making for vineyard management.

Computer Vision Synergy

Leverage advancements in deep learning for real-world applications.

Introducing U-Net

1

Convolutional Layers

Effectively extract spatial features from high-res images.

2

Encoder-Decoder

Capture global context and local detail for segmentation.

3

Skip Connections

Combine low-level and high-level features for precise outputs.

U-Net Architecture

1 **Convolutional Blocks**

Enhance feature extraction.

2 **Encoder Design**

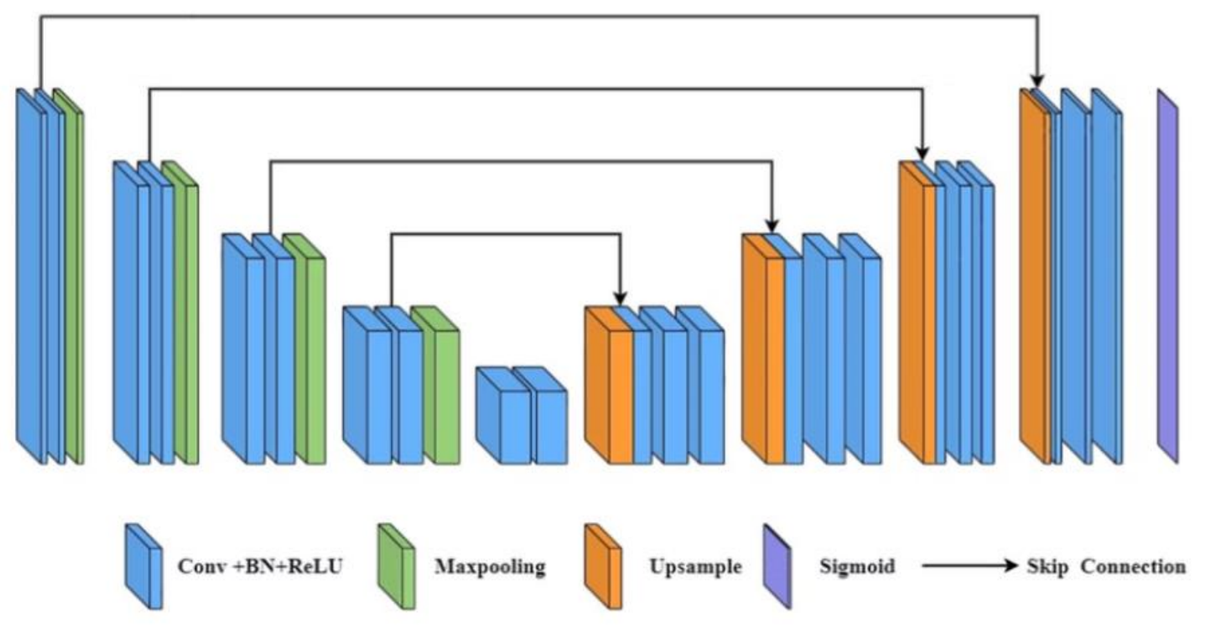
Convolution and max pooling for context.

3 **Bottleneck**

Intensive feature synthesis with 1024 filters.

4 **Decoder Design**

Upscales, retains details via skip connections.



Training U-Net

1

Data Augmentation

Techniques used - HorizontalFlipping, CoarseDropout, RandomBrightnessContrast

2

Loss Function and Callbacks

Loss Function - BinaryCrossEntropy

Callbacks - EarlyStopping, ModelCheckpoint, CSVLogger, ReduceLROnPlateau

3

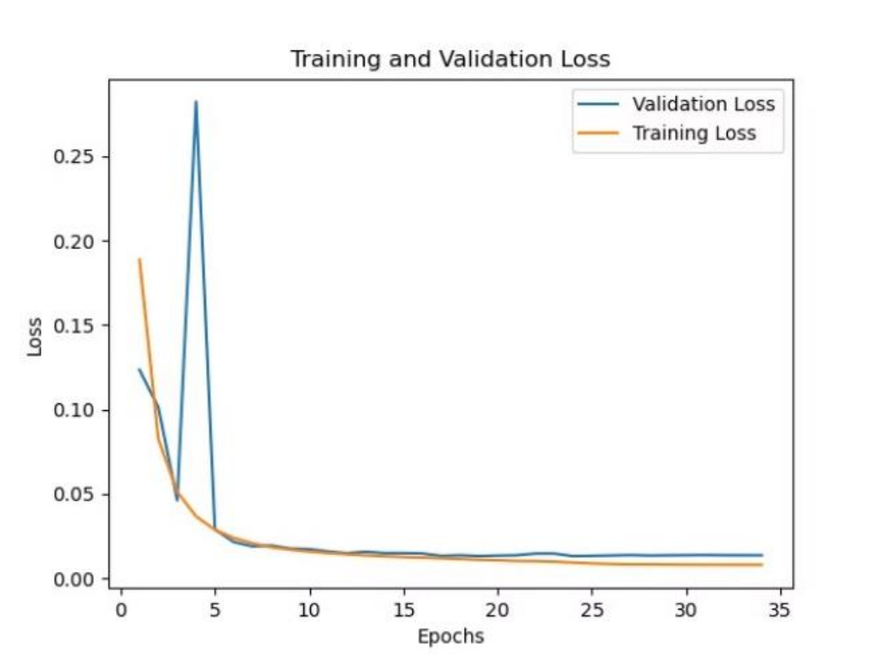
Hyperparameters

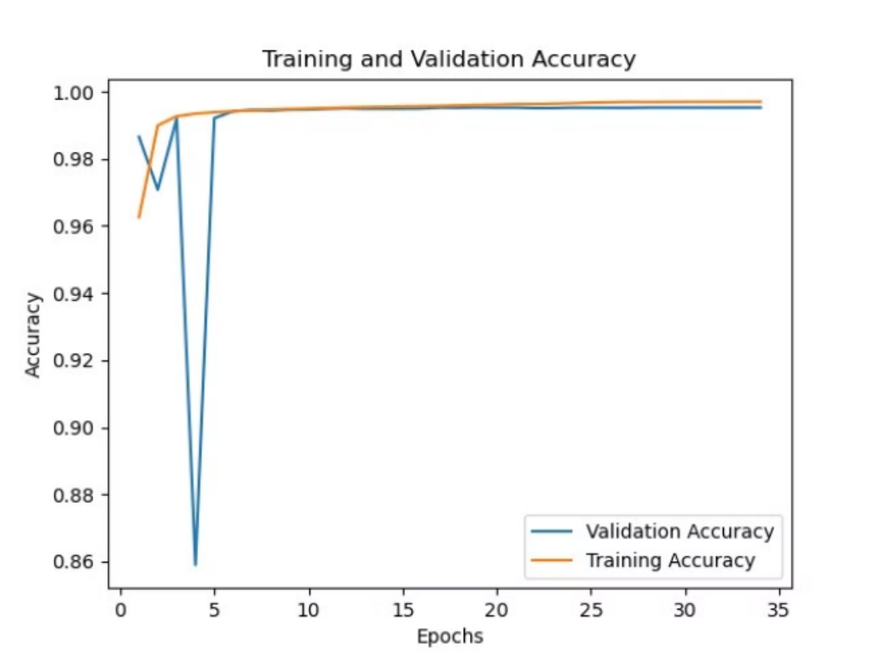
Epochs = 100, Learning Rate = 0.0001, Batch Size = 1,

Evaluating Performance

Metric	Description	Output
Mean IoU	Comprehensive metric for segmentation accuracy.	71.1%
Mean Inference	Time taken to process one image and output a result.	0.17664s (RTX 3090) approx.
Mean FPS	Number of images the model can process per second.	5.66122
Precision	Minimize false positives for reliable grape detection.	86.0%
Recall	Maximize true positives to capture all grapes.	81.1%

Plots





Output Images

Original image

Predicted Mask

Thank You

Open to Questions