YOLO Models for Instance Segmentation of Individual Tree Crowns from Aerial Imagery in Wellington

Ziyi Sun

Supervisors: Bing Xue and Mengjie Zhang

School of Engineering and Computer Science, Victoria University of Wellington

Background

Research Significance

- Prevent Landslide happened in the Wairarapa area
- Where exactly should we plant new trees?
- Which types of trees work best?

Landslide

Wairarapa Landscape

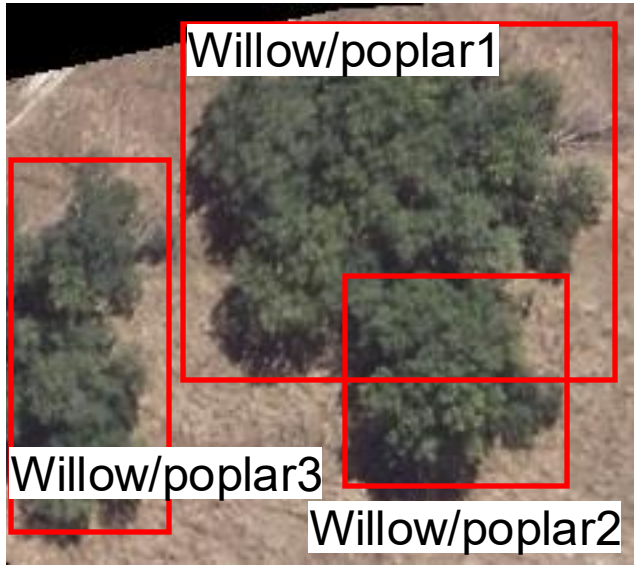
Background

What is Instance Segmentation of Individual Tree Crowns?

Goals:

- Classification: Assign a class label to each canopy in the image
- Detection: Provide the location of canopies in the form of bounding boxes
- Segmentation: Provide the shape of canopies inside each detected bounding box

(a) Image classification



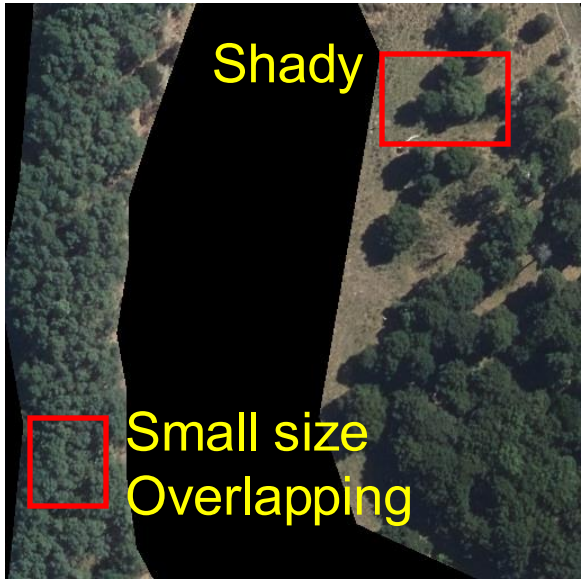
(b) Object detection

(c) Instance segmentation

Background

Challenges

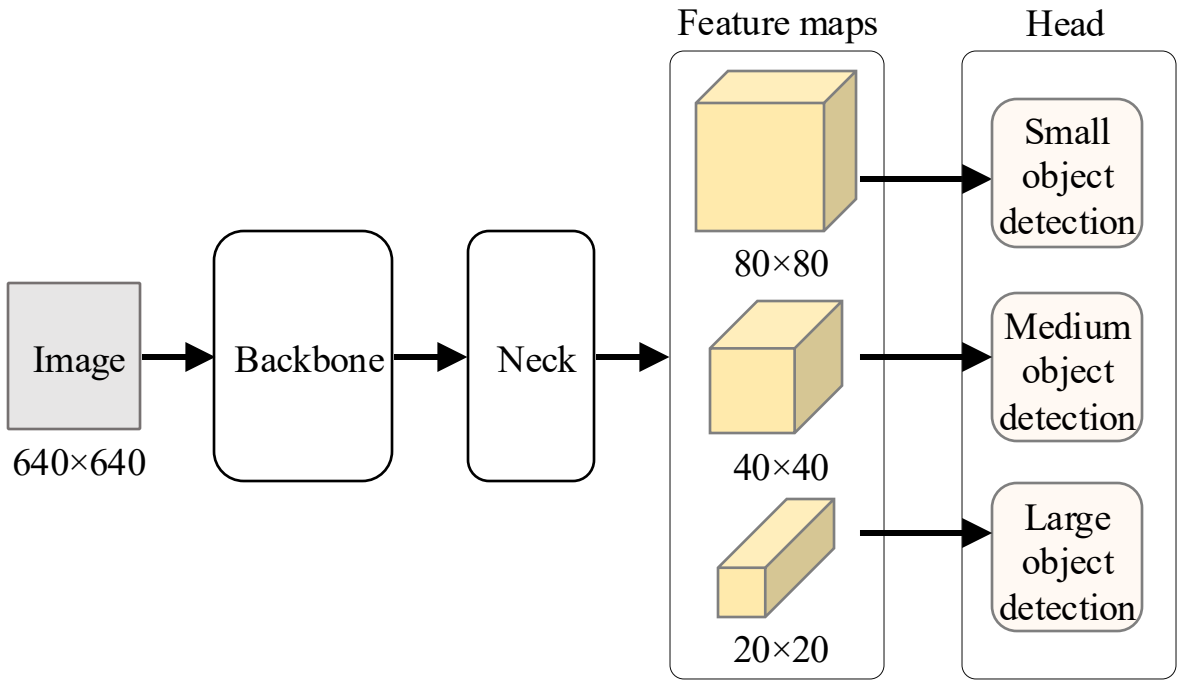
- Dense stands result in overlapping and shady canopies.
- Tree crown characteristics are complex, with variations in size, color, shape, and texture. Aerial imagery is difficult to provide sufficient information to distinguish variations in canopy characteristics, and there will be many small objects in the images.
- The background or surroundings are likely to be identified as false positives, such as shrubs, weeds, and grass.
- Uncertainties in data labelling



Background

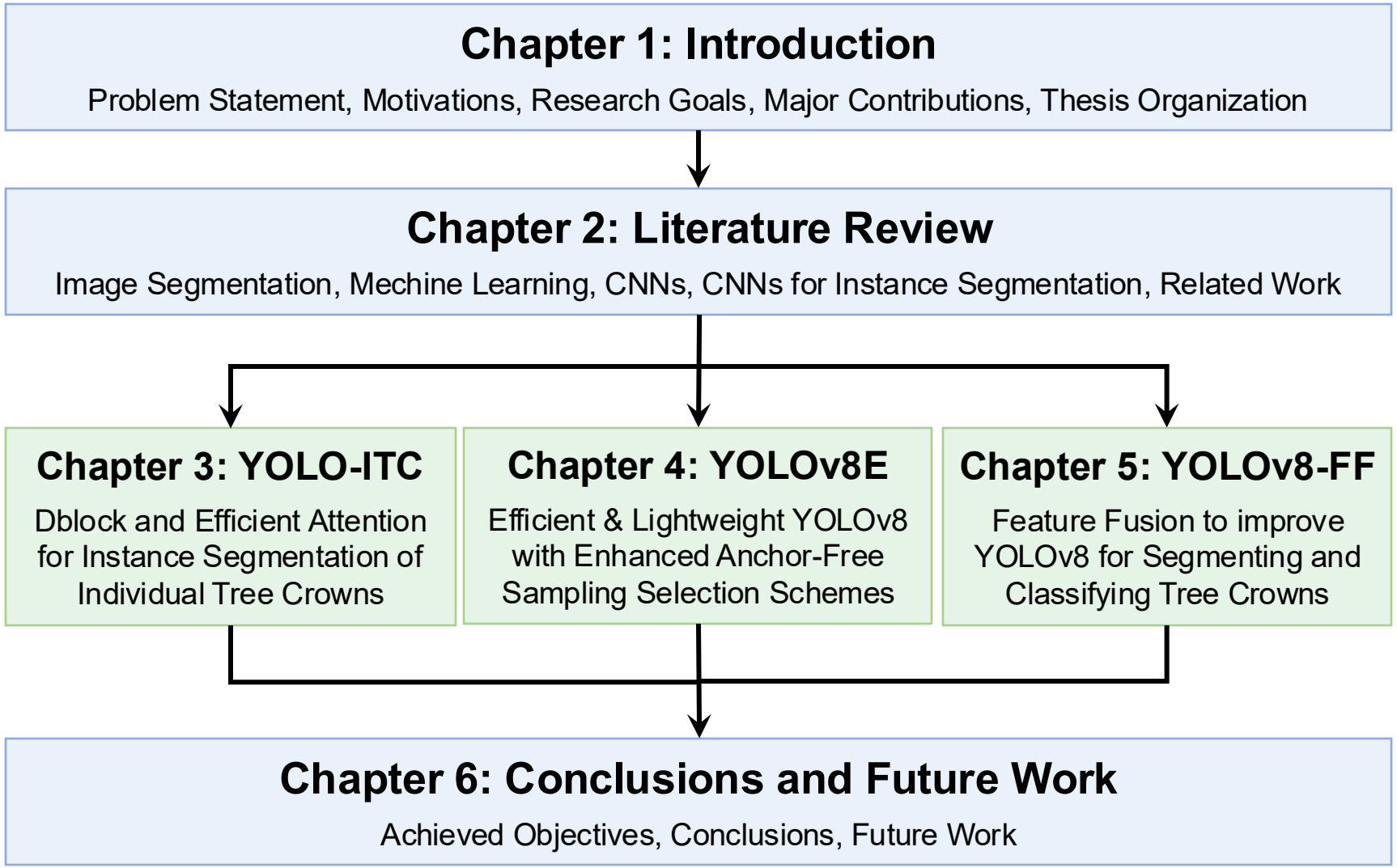
Why YOLO?

- Improved Accuracy and Speed
  - Efficient backbones and optimization algorithms
- Advanced Multi-Scale Detection
  - Handle diverse object sizes
- Robustness Against Varied Backgrounds
  - Enhanced feature extraction mechanisms and refined classification layers
- Real-time Processing Capabilities
  - The streamlined architectures and efficient processing techniques



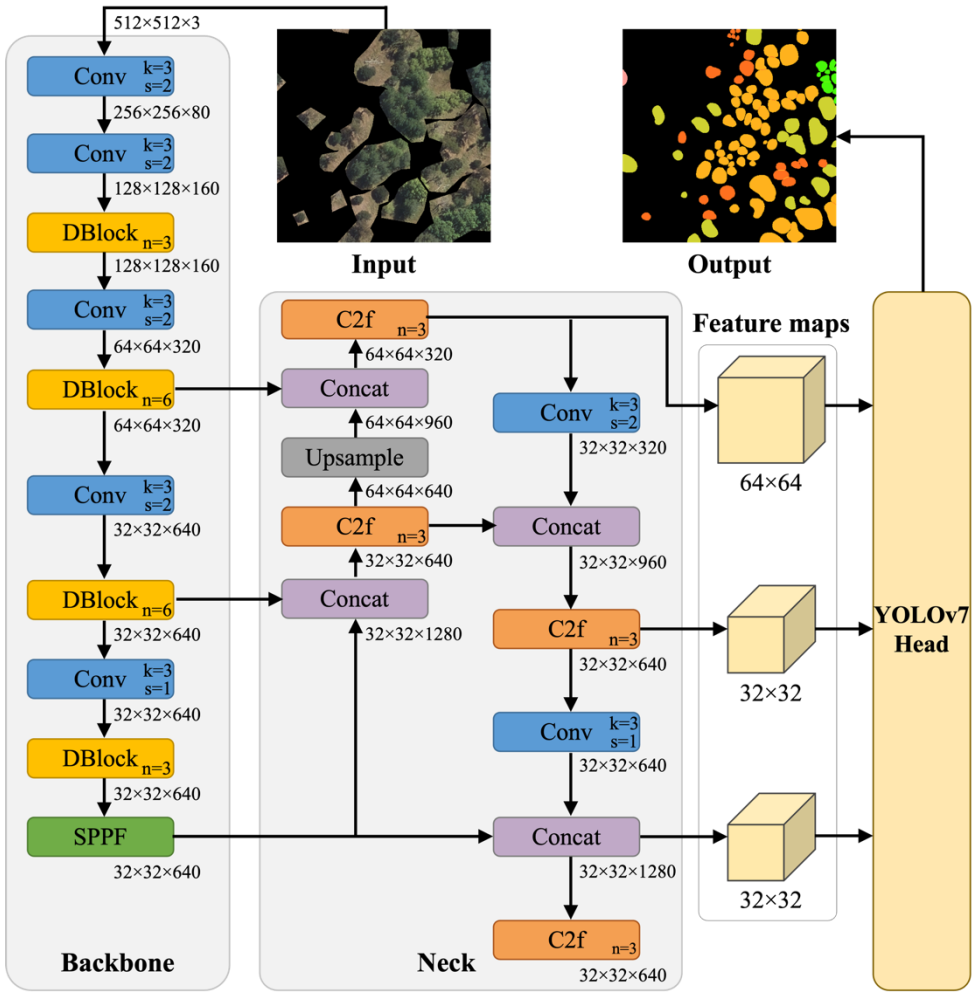
Background

Thesis Outline



Chapter 3: YOLO-ITC with Dense Block and Efficient Attention

Overall Framework

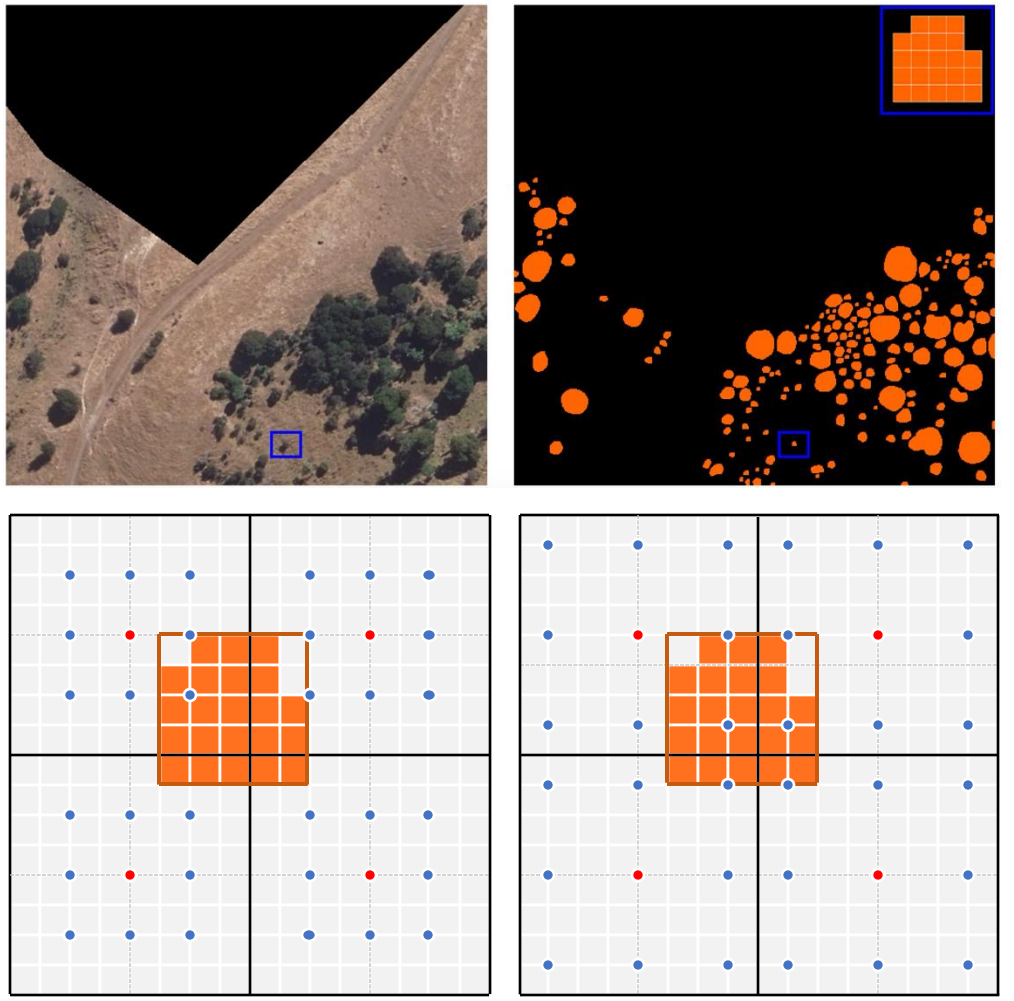


The goal of YOLO-ITC is to effectively address the instance segmentation of individual tree crowns in aerial imagery

- Demonstrate the strong effectiveness of YOLO-based architectures for the application of instance segmentation of individual tree crowns in Wellington
- Enhance feature extraction in dense, complex canopy environments by using Dense Blocks and Efficient Attention
- Reconfigure the large-object detection branch to focus on medium-scale canopies to enhance segmentation performance

Chapter 4: YOLOv8E: Efficient and Lightweight Model

Methodology



Two schemes for selecting candidate positive samples

The goal of YOLOv8E is to develop a more compact and accurate variant of the anchor-free YOLOv8 model.

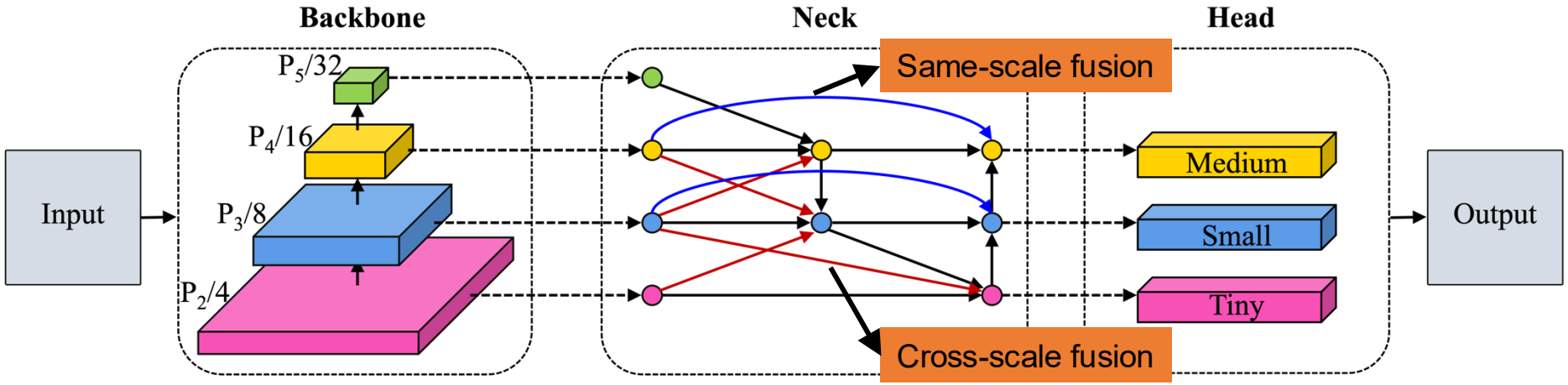
- Design two candidate positive sample selection schemes (A and B), by placing multiple keypoints within a grid cell, to raise the likelihood of capturing tiny canopies
- Efficient network design by omitting the large object detection branch, saving computational resources
- Adaptive hyperparameter tuning within the Task-Aligned Assigner

Chapter 5: YOLOv8-FF: Feature Fusion-Enhanced Model

Methodology

The goal of YOLOv8-FF is to propose a feature fusion-enhanced model based on YOLOv8 to address diverse canopy sizes

- Design feature fusion techniques that incorporating same-scale and cross-scale fusion to preserve fine-grained details while capturing global context
- Propose a refined network design for identifying tiny, small, and medium crowns
- Incorporate large-kernel convolutions to expand the receptive field, improving the identification of medium-sized canopies



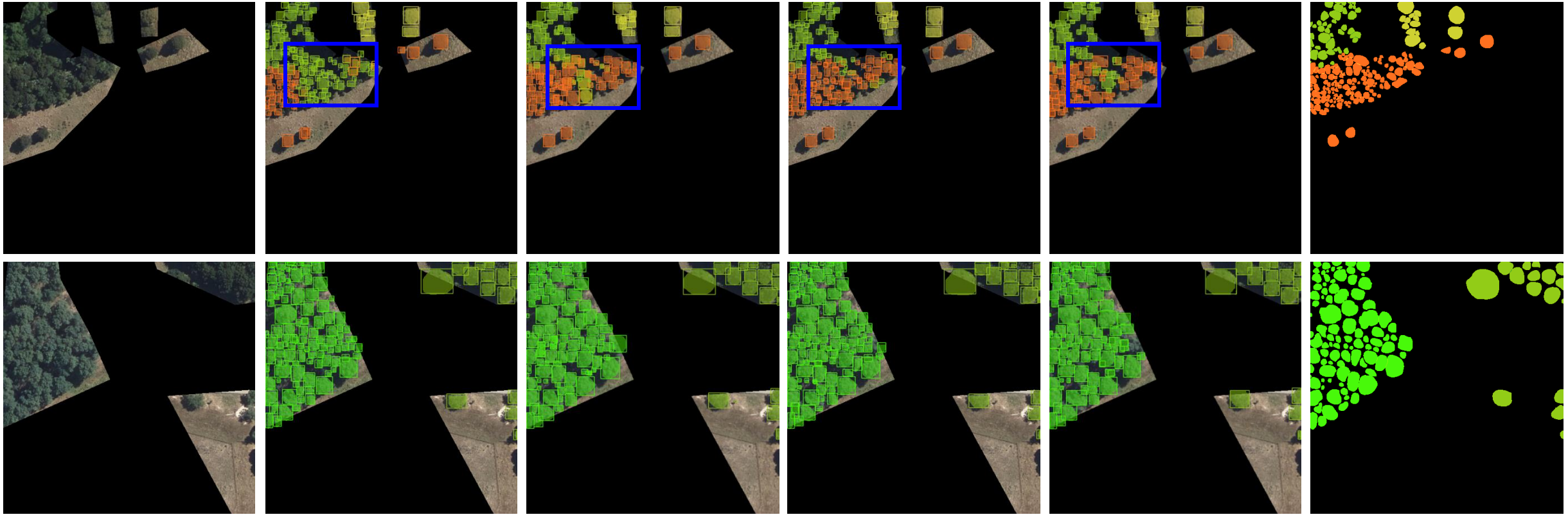
Results

Quantitative Results

Method	Box AP	Mask AP	AP_{50}	AP_{75}	AP_S	AP_M	FLOPs	#Params
YOLO-ITC	36.2	32.0	60.3	31.7	30.8	45.3	141.8G	82.2M
YOLOv8E-A	36.0	32.1	60.7	32.2	30.9	**45.4**	106.4G	56.9M
YOLOv8E-B	36.1	32.2	60.5	31.9	31.2	43.7	106.4G	56.9M
YOLOv8-FF	**37.6**	**32.9**	**61.4**	**32.8**	**31.9**	44.6	196.1G	54.4M

Results

Visual Results



Contributions

This thesis develops innovative YOLO models to perform instance segmentation and species classification of individual tree crowns from aerial imagery by enhancing the feature extraction network, refining the detection framework, and developing efficient network architectures.

- Develop **YOLO-ITC** with a new detection scheme, dense blocks with efficient attention to enhance the effectiveness of YOLO-based architecture.
- Propose **YOLOv8E** with two novel sample selection strategies to better handle small canopies and the redesigned network to improve accuracy while reducing the model size and parameter count.
- Propose **YOLOv8-FF**, a feature fusion-enhanced model that integrates cross-scale and same-scale fusion, alongside a sparse large kernel network for better global context aggregation.

Future Work

- Overlapping and Ambiguously Bounded Tree Crowns
- Model Generalization
- Multi-Sensor and Multi-Spectral Data
- Transformer-Based Architectures
- Semi-Supervised and Active Learning Approaches
- Real-Time and On-Device Implementation
- Neural Architecture Search

Publications

[1] **Ziyi Sun**, Bing Xue, Mengjie Zhang, and Jan Schindler. “Feature Fusion to Improve YOLOv8 for Segmenting and Classifying Aerial Images of Tree Crowns.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2025): 1-14.

[2] **Ziyi Sun**, Bing Xue, Mengjie Zhang, and Jan Schindler. “YOLO-ITC: A New YOLO Method for Instance Segmentation of Individual Tree Crowns.” *IEEE Transactions on Emerging Topics in Computational Intelligence* (2025): 1-15.

[3] **Ziyi Sun**, Bing Xue, Mengjie Zhang, and Jan Schindler. “YOLOv8E: an efficient YOLOv8 method for instance segmentation of individual tree crowns in Wellington City, New Zealand.” *Journal of the Royal Society of New Zealand* (2024): 1-26.

[4] **Ziyi Sun**, Bing Xue, Mengjie Zhang, and Jan Schindler. “An Improved Mask R-CNN for Instance Segmentation of Tree Crowns in Aerial Imagery.” *2023 38th International Conference on Image and Vision Computing New Zealand (IVCNZ),* Palmerston North, New Zealand, 29-30th November 2023. pp. 1-6.

Thank you !