



TIMBERVISION: A MULTI-TASK DATASET AND FRAMEWORK FOR LOG-COMPONENT SEGMENTATION AND TRACKING IN AUTONOMOUS FORESTRY OPERATIONS

Dataset Design

Creating a large-scale multi-task **dataset** and efficient semi-automatic **annotation pipeline** for **trunk-component** detection and tracking

Image Acquisition

- Affordable **RGB** sensors
- Multiple application **scenarios**
- Diverse **locations** and **environmental conditions**



Log-Component Annotation

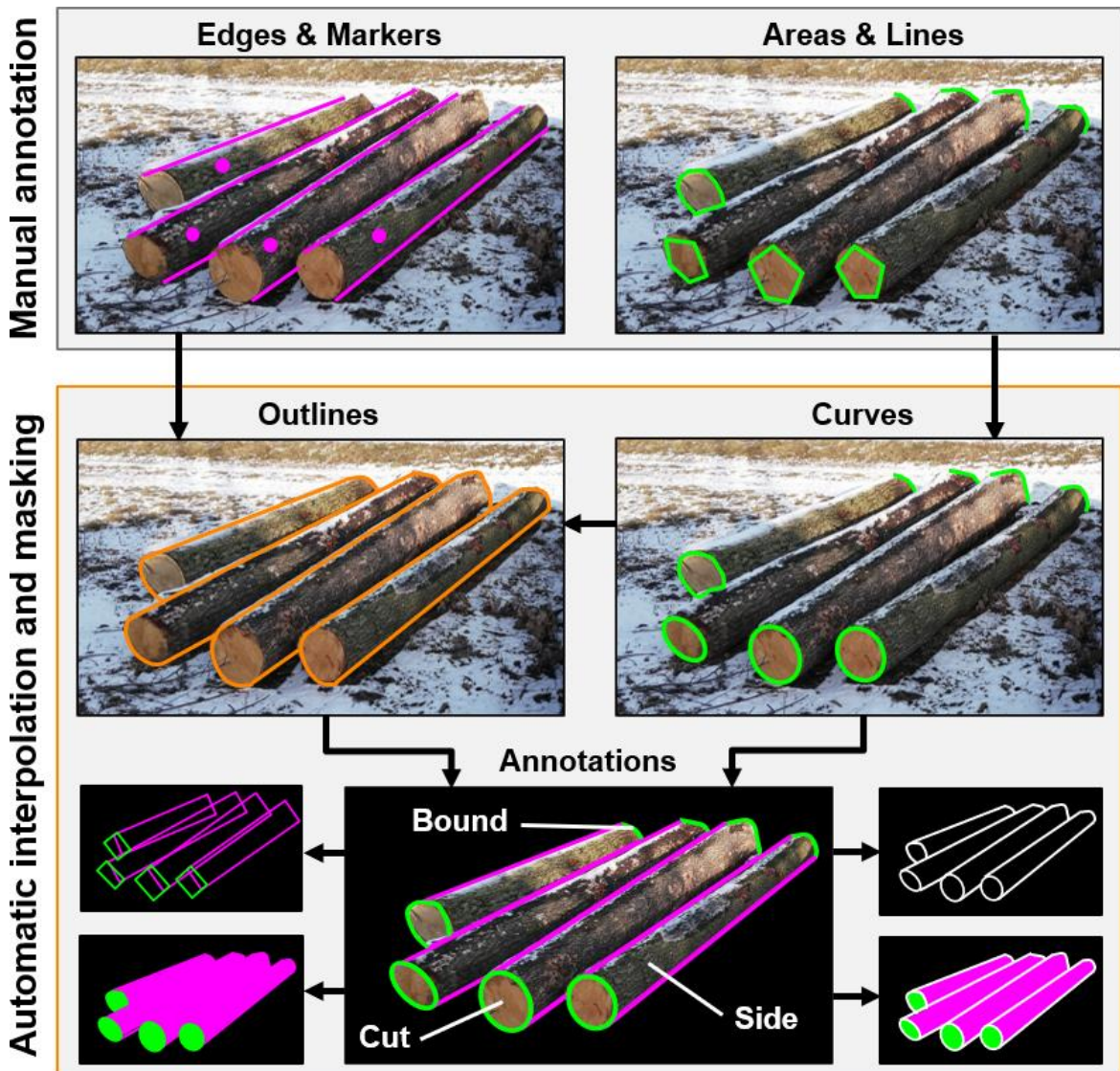
- Efficient manual **delineation**
- Automatic inference** of multi-task **Cut**, **Side** and **Bound** annotations
- Largest** trunk-segmentation dataset in terms of **real images** and **annotation detail**

Subset	Rec	Images	Trunks	Comp
Core	71	1 415	17 999	33 445
Loading	18	220	3 853	8 782
Harvesting	10	45	825	1 720
OpenSource	42	42	776	1 634
Tracking	13	266	1 900	3 708
TimberSeg*	10	35	940	2 049
	164	2 023	26 293	51 338

Table 1. TimberVision dataset statistics. Images are clustered into **Recording** sessions based on timestamps and GPS tags, ensuring a difference of at least two minutes and one hundred meters between them. Trunks are divided into their three **Component** classes.

Scene-Parameter Annotation

- Object** properties and **recording context**
- Quantifying **data variability** and **sampling** images

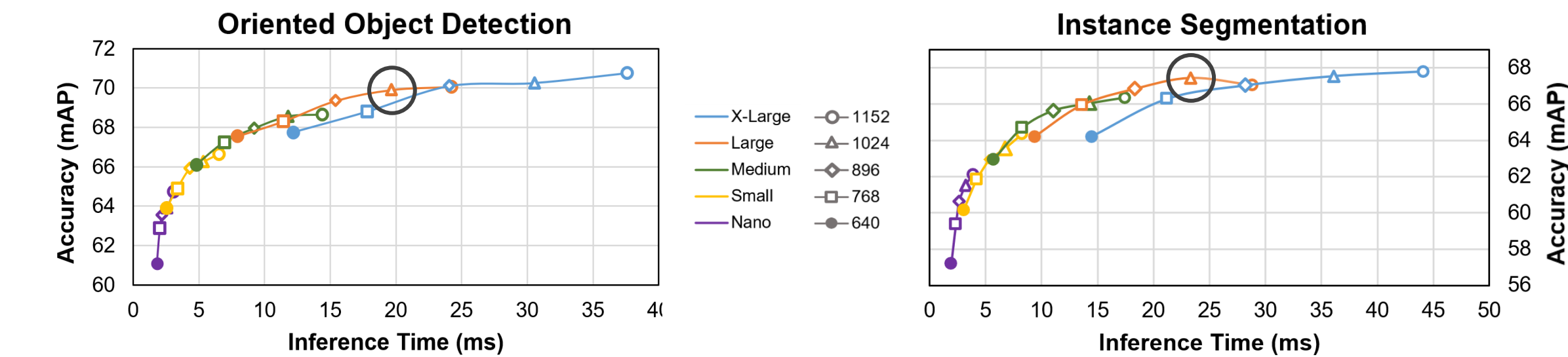


Experimental Evaluation

Combining **oriented object detection** and **instance segmentation** to robustly detect tree trunks under **challenging conditions**

Model Ablation

- Evaluating multiple YOLOv8 **model capacities** and **images sizes**



Performance Evaluation

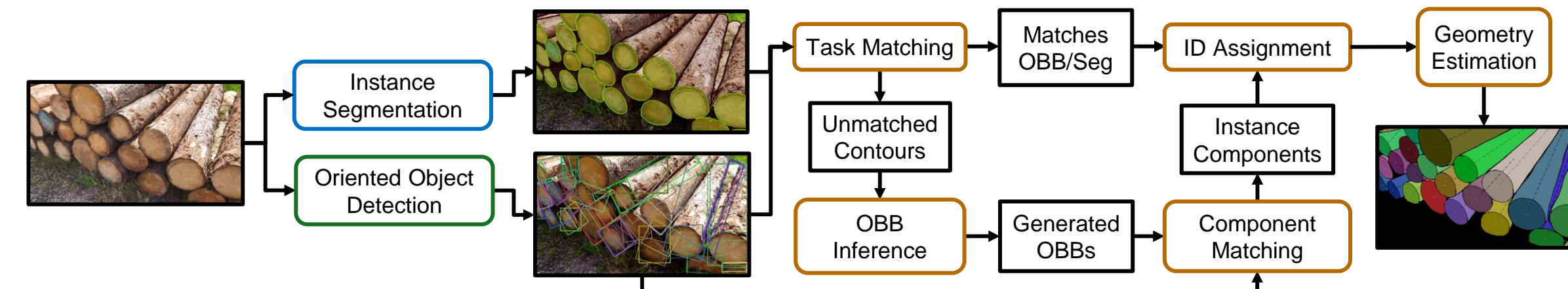
- By **scenario** and **component**
- Influence of **scene parameters**

		Core	Load	Harvest	Open
OOD	Cut	83.0	79.5	80.4	80.8
	Side	59.8	66.3	47.1	45.0
ISEG	Cut	77.4	67.0	67.7	72.1
	Side	61.2	69.7	45.3	45.4

Table 4. Model performance as mAP⁵⁰⁻⁹⁵ for the classes *Cut* and *Side* for test and validation images of the **Core**, **Loading**, **Harvesting** and **OpenSource** subsets.

Task Fusion

- Combining results for higher **robustness** and **geometric cues**



Multi-Object Tracking

Preserving **object identities** over time and **smoothing results** by tracking fused trunk instances

Performance Evaluation

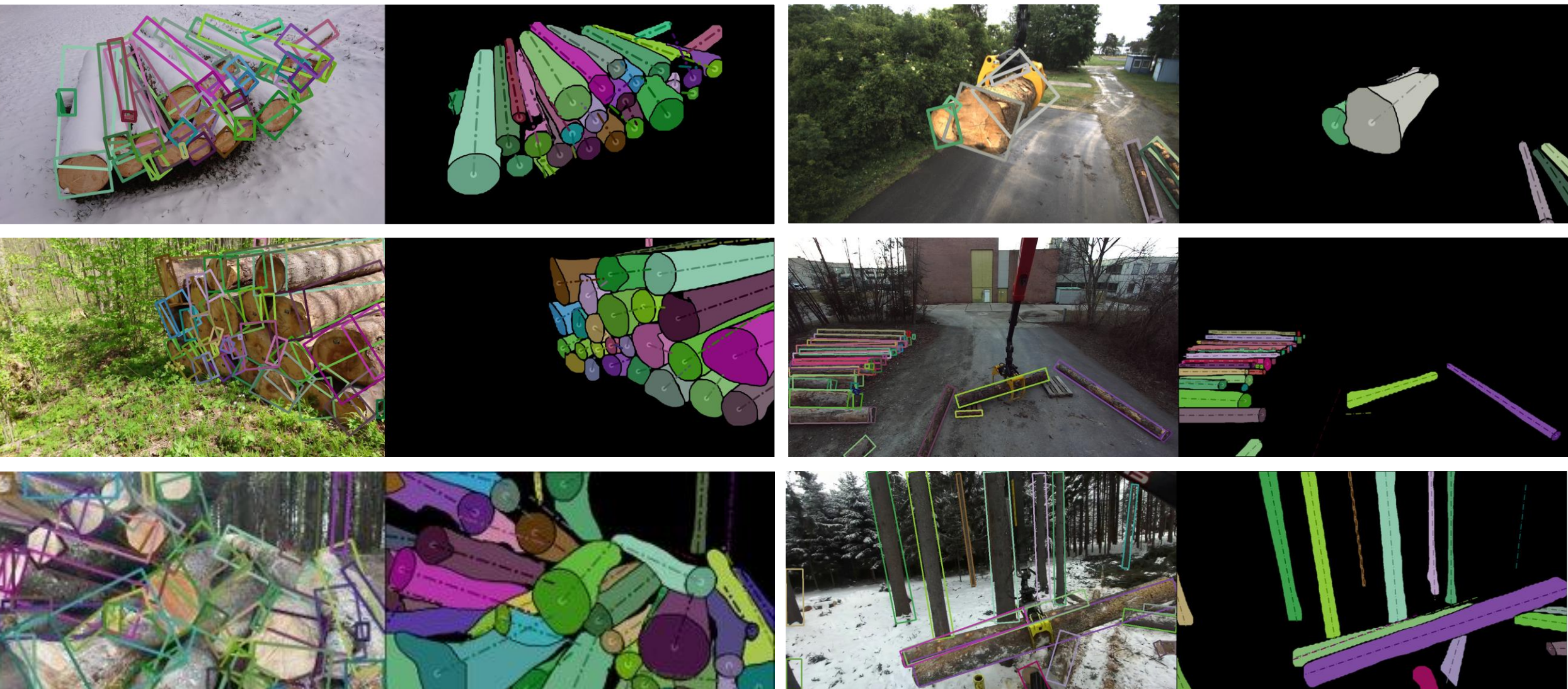
- Evaluation on dedicated **subset** of videos at multiple **framerates**
- Multiple **configurations** of two tracking approaches

	IDF1	IDP	IDR	MOTA	mIoU _c
ByteTrack	71.1	85.0	61.2	63.8	85.8
Bot-SORT	72.5	86.6	62.4	63.7	85.4
Optimized	72.9	86.0	63.2	65.2	85.1
Opt 10 fps	66.5	81.1	56.4	57.8	84.4

Table 4. Mean MOT results for default and optimized configurations on all *Tracking* sequences at a default frame rate of 30 fps.

Conclusion

Robust **trunk segmentation** and **tracking** performance across most scenarios as wells as external public image data



Outlook

- Hybrid model** for oriented object detection and instance segmentation
- Combination of monocular approach with **depth sensors**