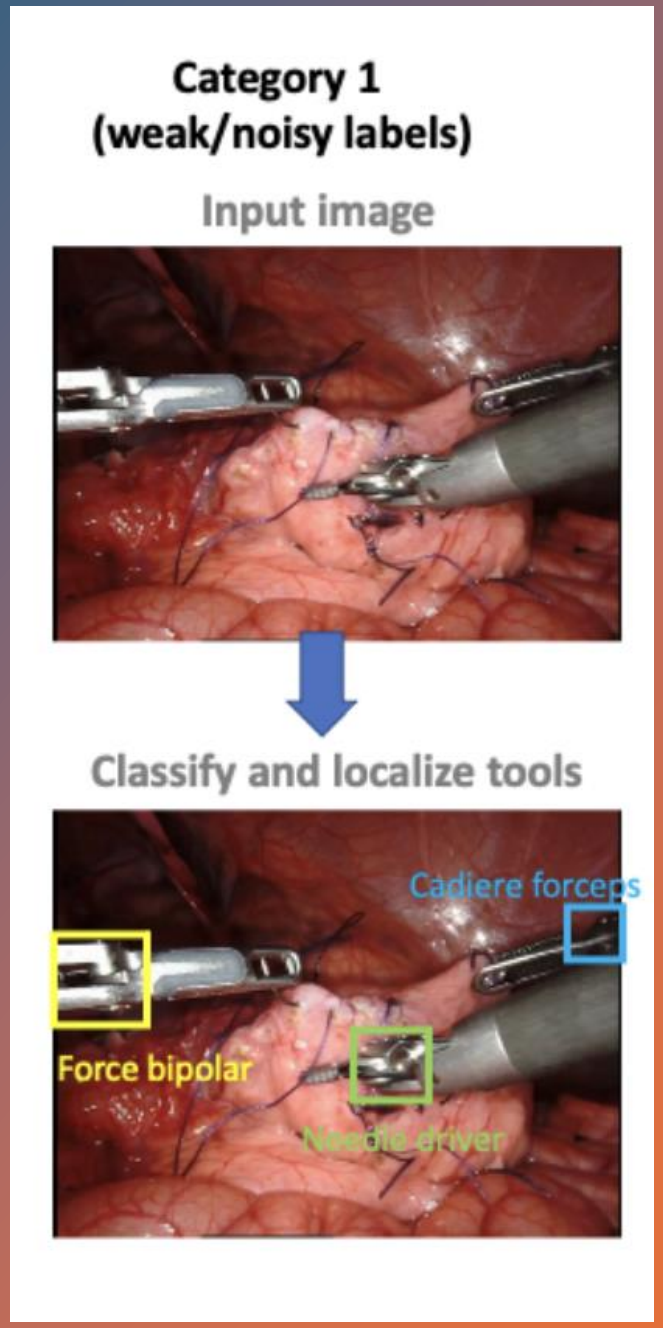
SURGVU 2026 · WEEKLY PROGRESS REVIEW

Week 4 Presentation

DeAndre Bailey

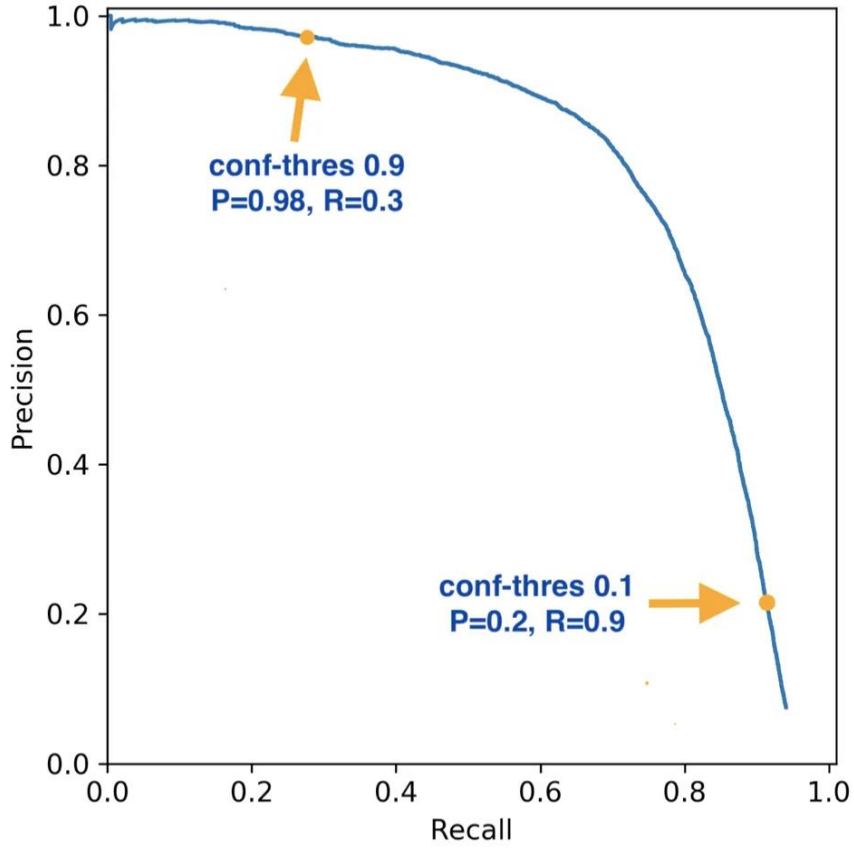


Project Recap

- SurgVU Challenge
- Category 1 Classification and Localization
- 12 official tool classes that must be detected
- mAP@50 is the primary leaderboard score

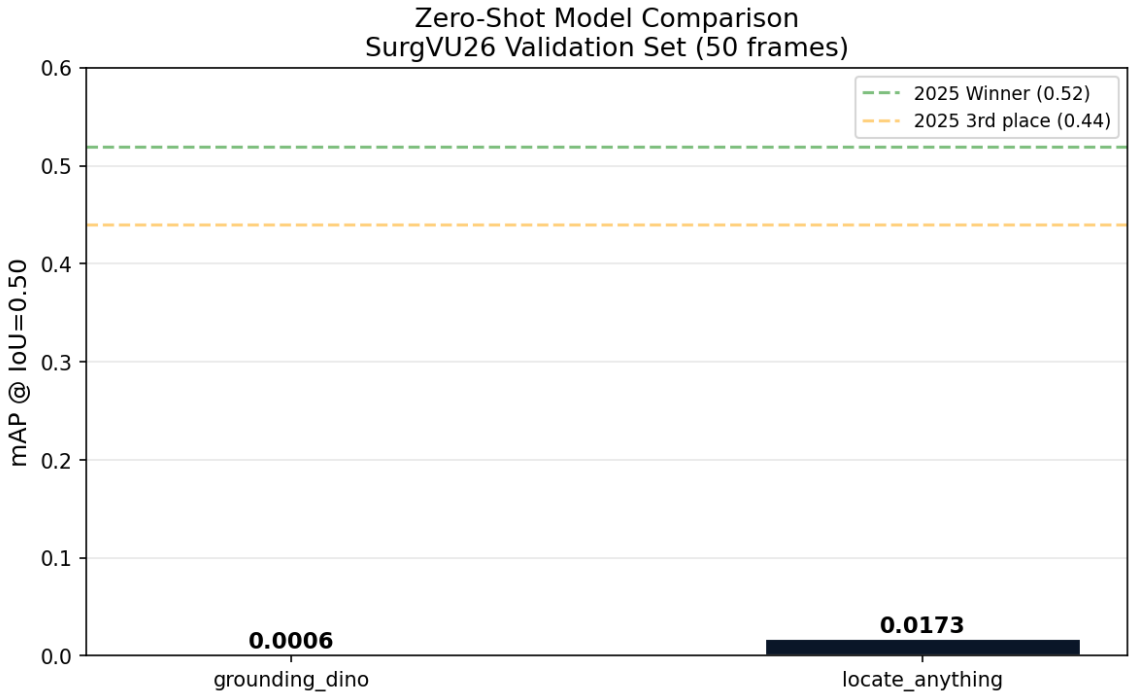
mAP@50

- mAP@50 = mean Average Precision at IoU threshold 0.50
- IoU (Intersection over Union)
  - Precision = TP / TOTAL PREDICTIONS
  - Recall = TP / TOTAL GROUND TRUTHS
- Average Precision (AP)
  - Precision Recall Curve
- Mean Average Precision



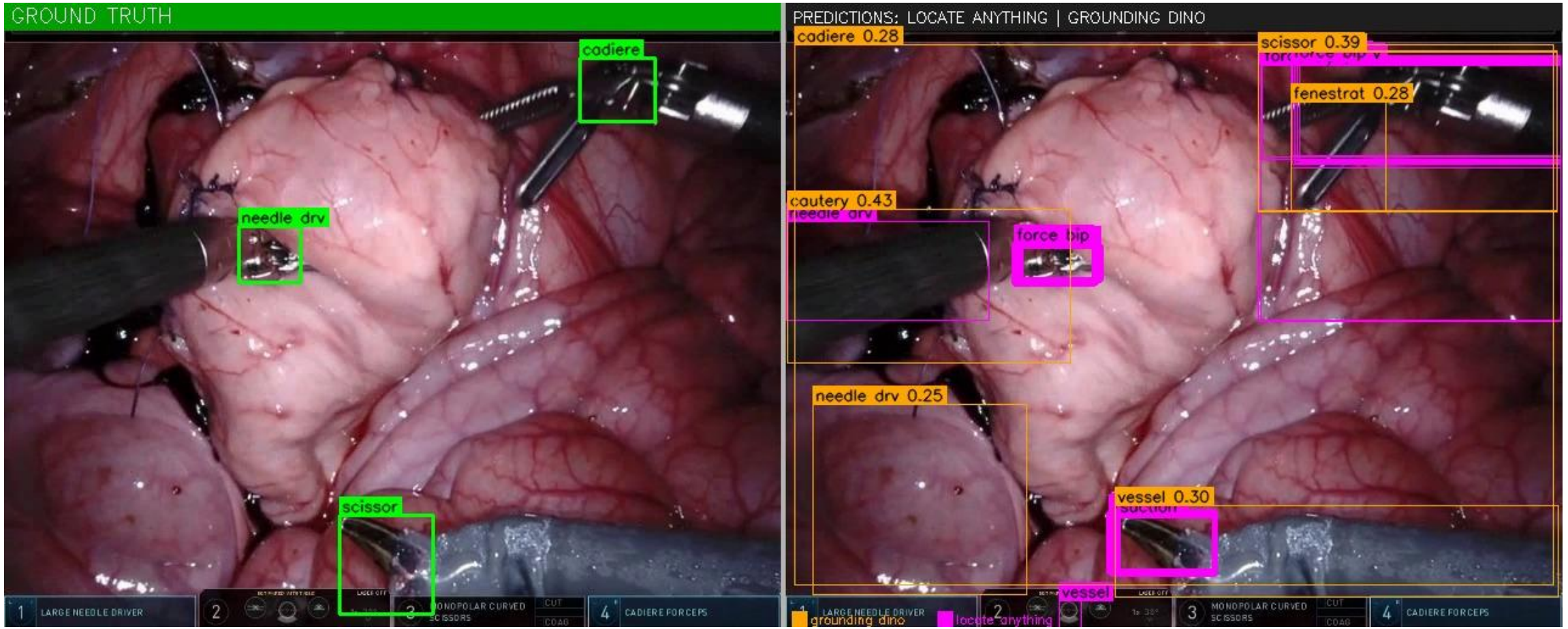
Baseline (Model Comparison)

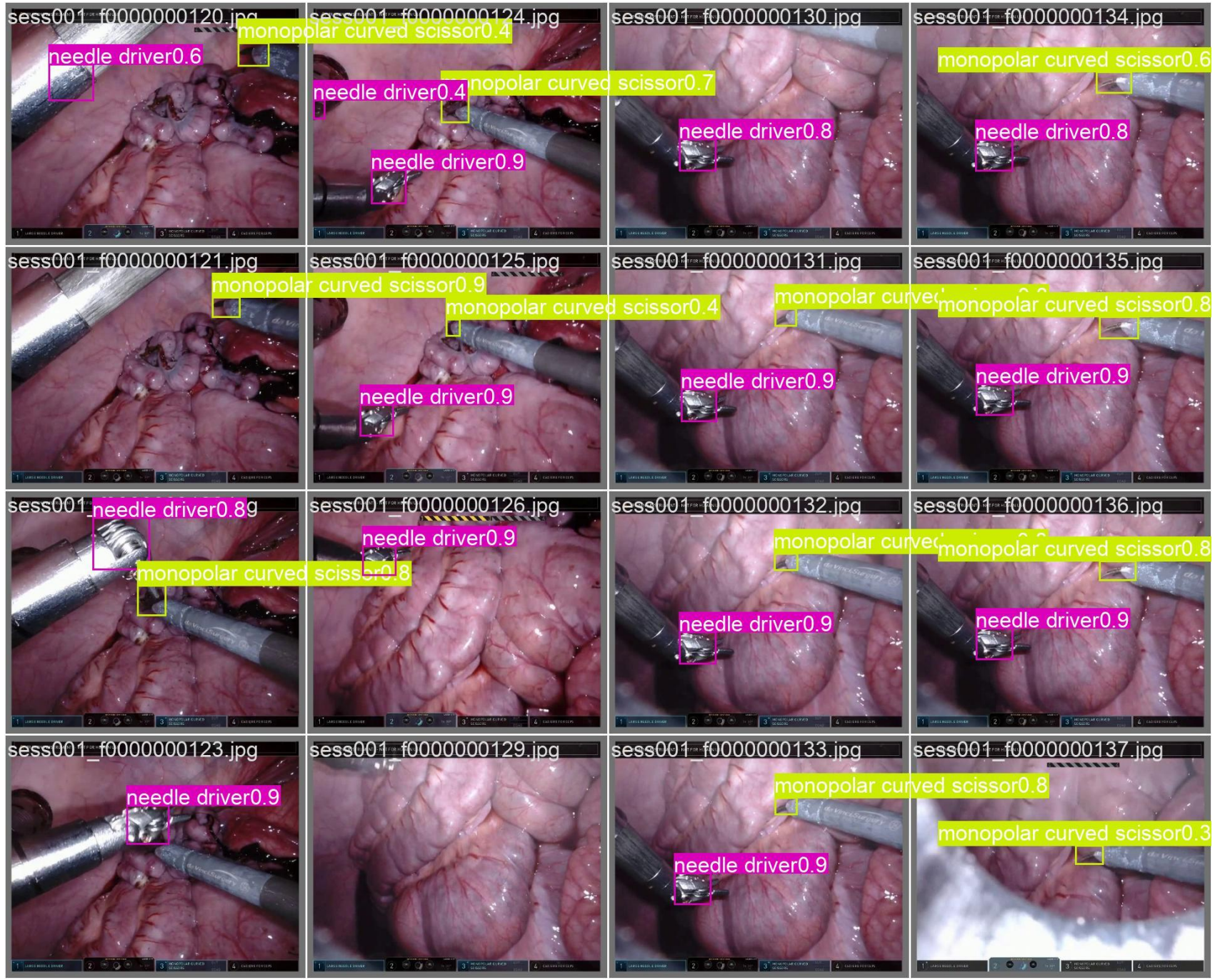
- Zero-Shot Models vs Winners
- Table showing mAP@50 evaluation on Validation set split



Model	Type	mAP@50
Grounding DINO (zero-shot)	Zero-shot	0.0006
LocateAnything-3B (zero-shot)	Zero-shot	0.0173
YOLOv8n (fine-tuned, 50 epochs)	Fine-tuned	**0.553**
YOLOv8l (fine-tuned, 50 epochs)	Fine-tuned	0.531
SurgVU 2025 Winner (reference)	Fine-tuned	0.519

Baseline (Visual Comparison)




sess001_f0000000120.jpg
monopolar curved scissor0.4
needle driver0.6
sess001_f0000000124.jpg
needle driver0.4
nopolar curved scissor0.7
needle driver0.9
sess001_f0000000130.jpg
needle driver0.8
sess001_f0000000134.jpg
monopolar curved scissor0.6
needle driver0.8
sess001_f0000000121.jpg
monopolar curved scissor0.9
sess001_f0000000125.jpg
monopolar curved scissor0.4
needle driver0.9
sess001_f0000000131.jpg
needle driver0.9
sess001_f0000000135.jpg
monopolar curved scissor0.8
needle driver0.9
sess001_needle driver0.8g
monopolar curved scissor0.8
sess001_f0000000126.jpg
needle driver0.9
sess001_f0000000132.jpg
needle driver0.9
sess001_f0000000136.jpg
monopolar curved scissor0.8
needle driver0.9
sess001_f0000000123.jpg
needle driver0.9
sess001_f0000000129.jpg
sess001_f0000000133.jpg
monopolar curved scissor0.8
needle driver0.9
sess001_f0000000137.jpg
monopolar curved scissor0.3

Our Labeled Data Covers 8 of 12 Tool Class

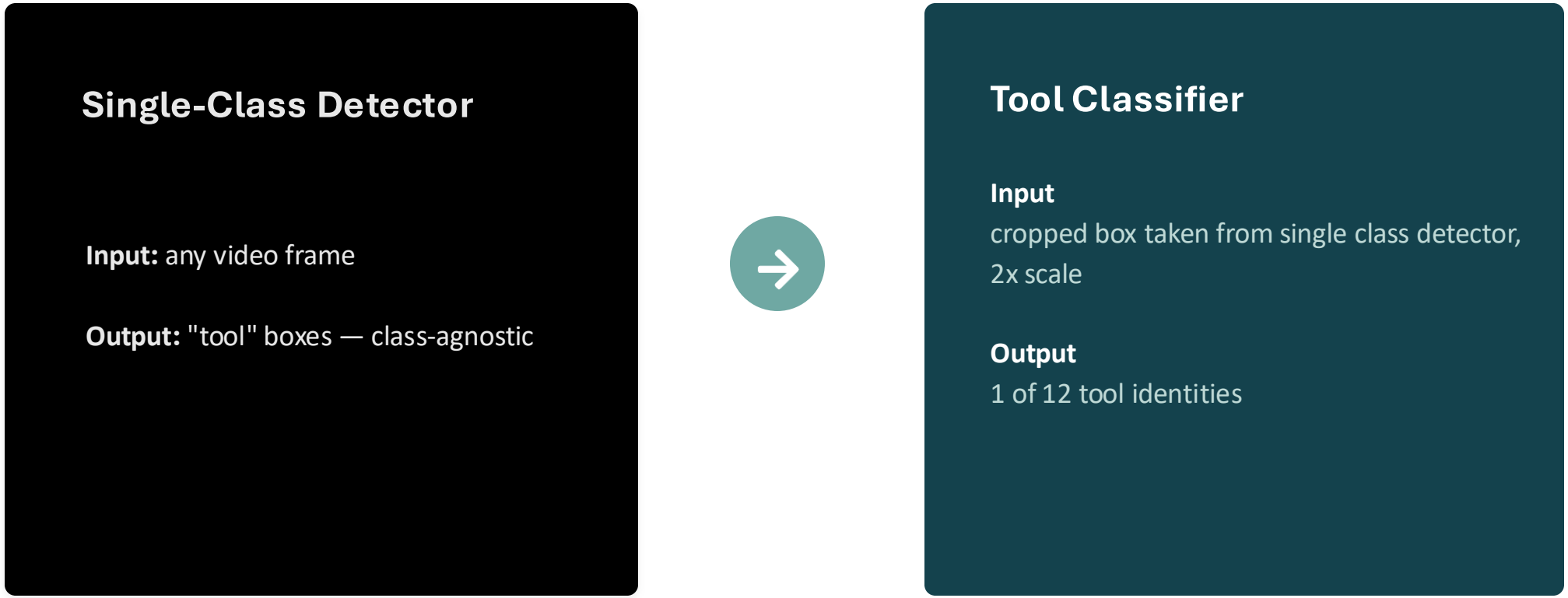
Tool	Val Boxes
Needle driver	3,584
Monopolar curved scissor	2,833
Cadiere forceps	1,871
Bipolar forceps	1,798
Force bipolar	781
Clip applier	258
Grasping retractor	185
Vessel sealer	13

- 4 classes with zero ground-truth boxes in the validation set
- A model can't pseudo-label a tool it has never been trained to recognize.

Missing Stapler, Permanent cautery hook/spatula, Tip-up fenestrated grasper, Prograsp forceps

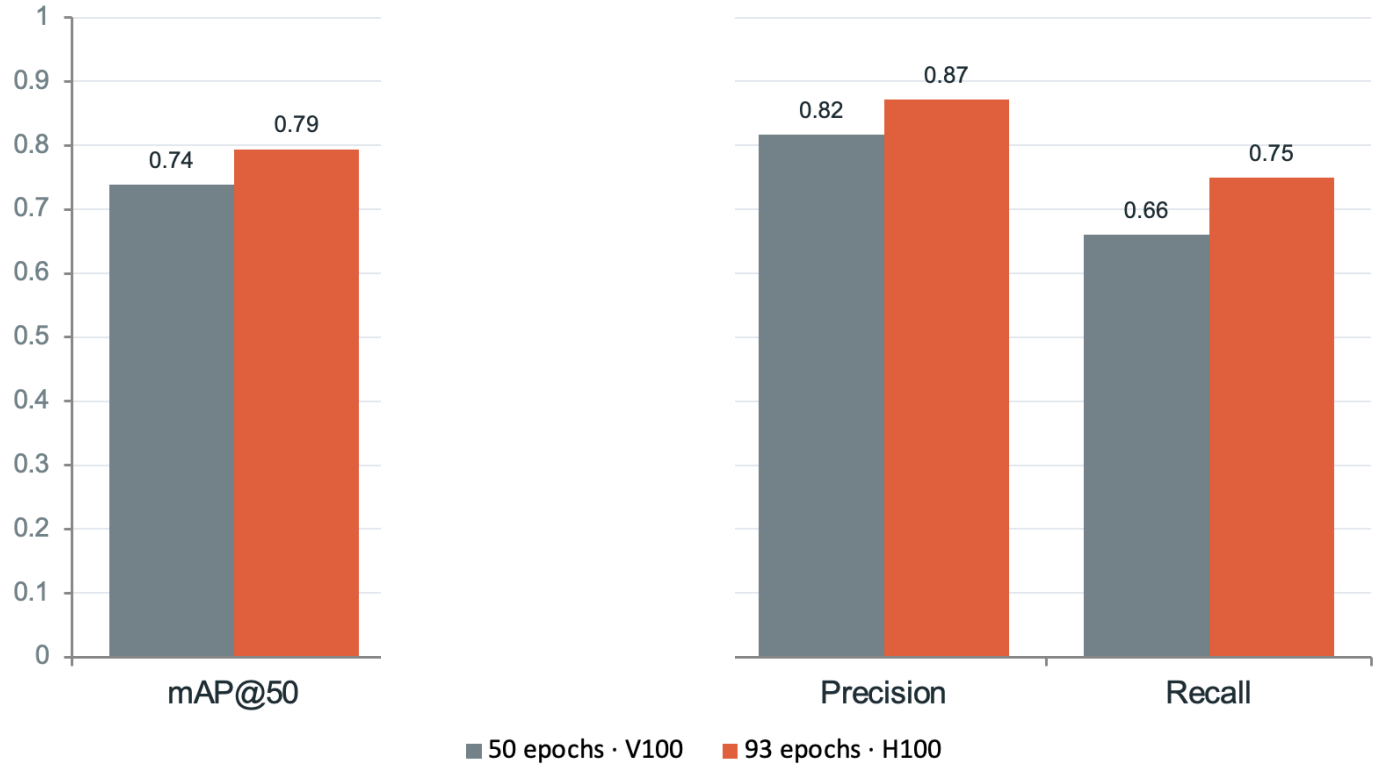
Methodology

- Rather than detecting 12 tool classes at once, split the problem
- Two Stage Pipeline Inspired by Medibot (2025)



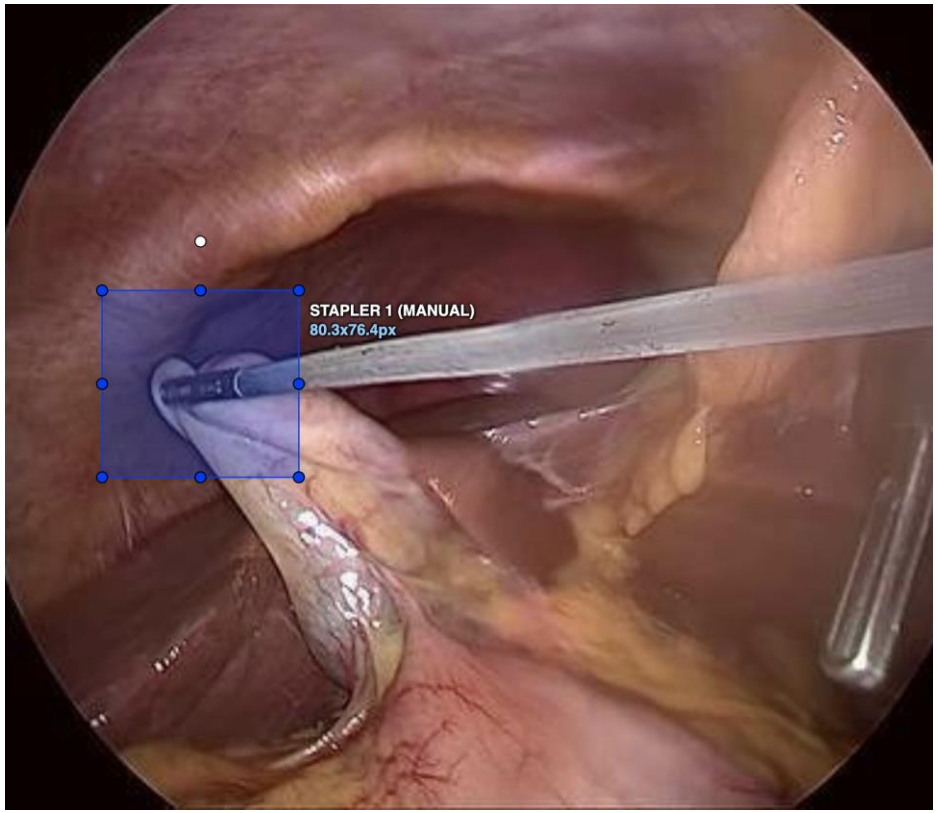
Single Class detector training

- Recall is the most important
- Using 200 epochs (early stopped at 93) to train the model Instead of 50
- the 0.66 → 0.749 Recall jump



Getting Ready for Manual Annotation

- Find ~100-150 frames per class, ~500 total
- Frames sampled from the middle of each tool-install period
- Setup CVAT annotation tool



Example: stapler annotated at the tip, 80.3 × 76.4px — case_030

Next Steps

- Receive the 155 Sessions (SurgVU Challenge Training Data)
- Run the single detector
  - Generate tool boxes
- Manual Annotate frames
  - Additionally find more Datasets that fit
- Begin training classifier
  - Find the best model for classifying

Thank You!