

Efficient Annotation of Segmentation Datasets with Polygon-RNN++

Sanja Fidler

Presented by Amlan Kar

University of Toronto, Vector Institute

Toronto



Team



Luis
Castrejon



Kaustav
Kundu



Raquel
Urtasun



David
Acuna



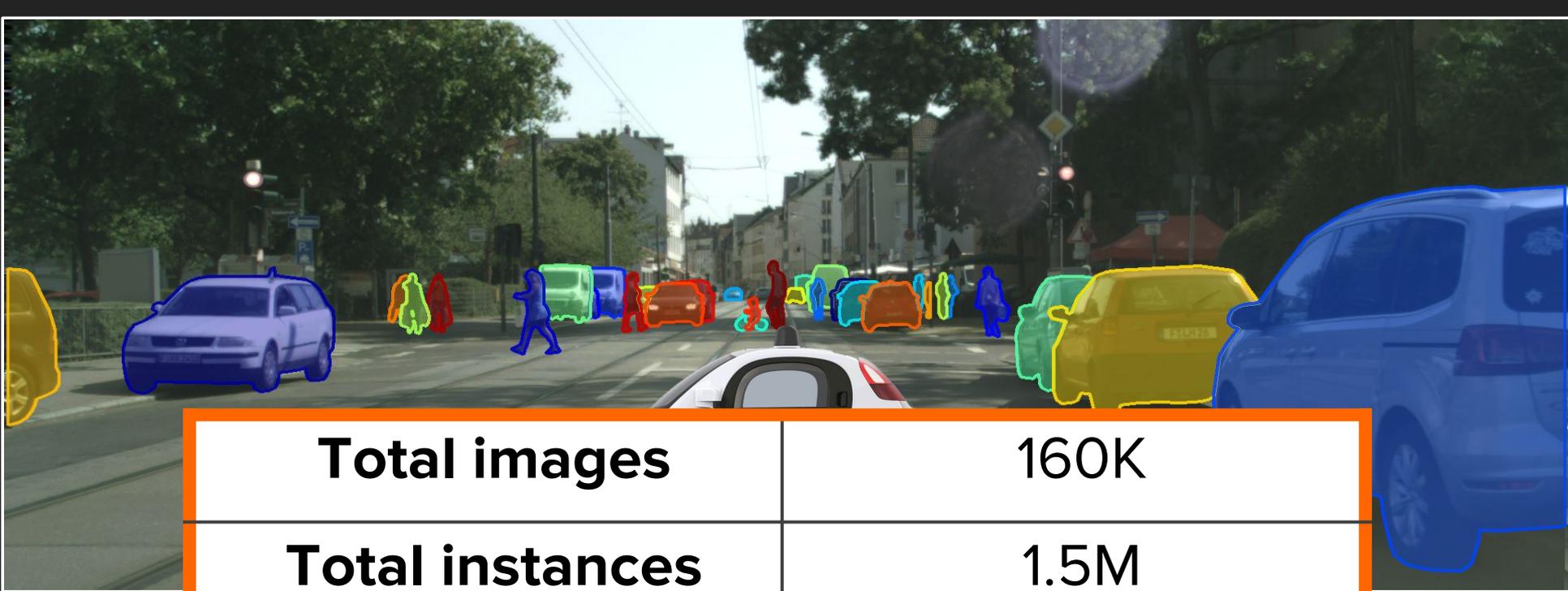
Huan Ling



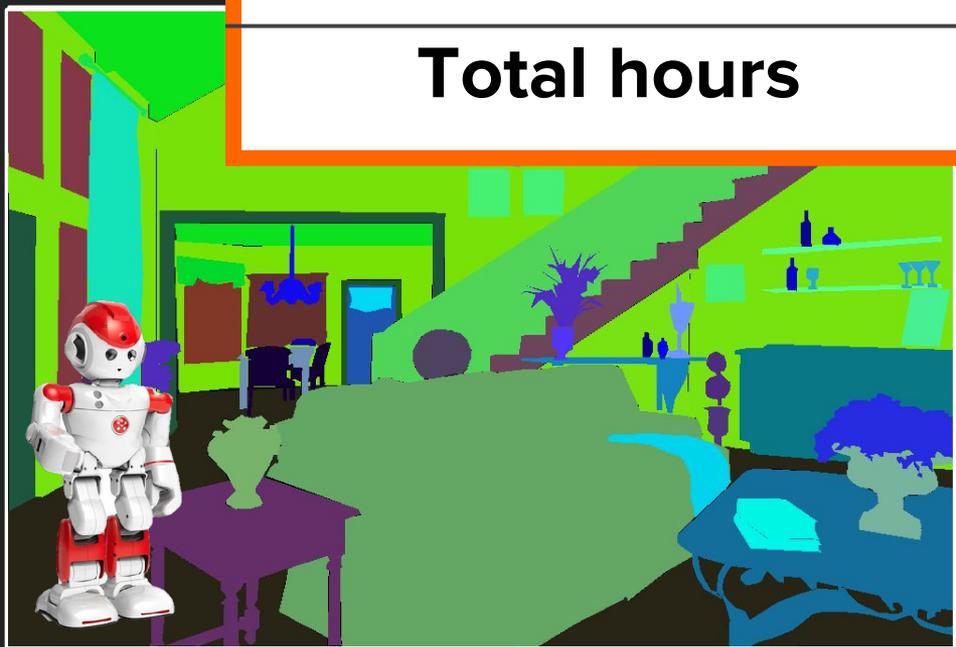
Amlan Kar

Castrejon, Kundu, Urtasun, Fidler, Annotating Object Instances with PolygonRNN. CVPR'17

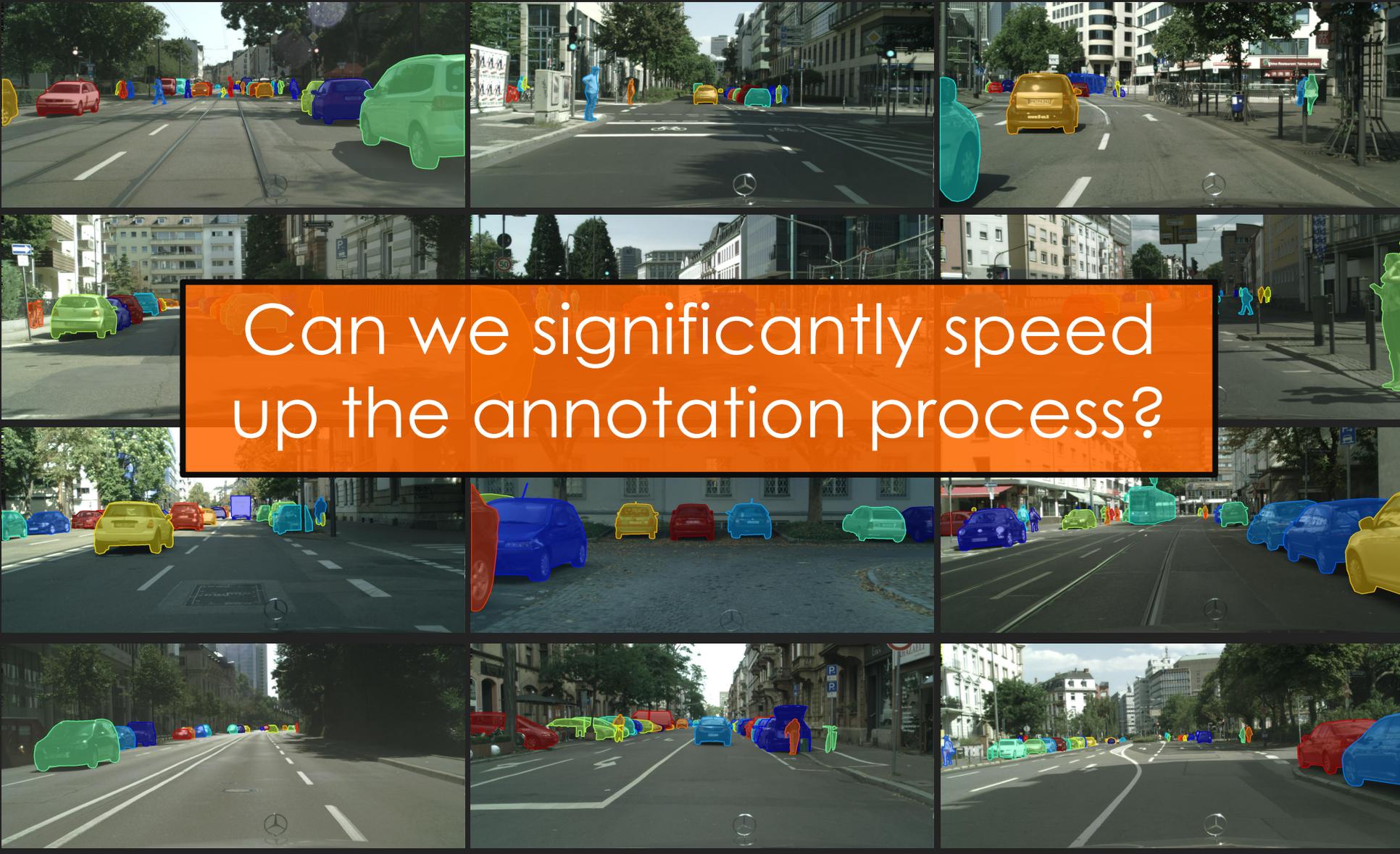
Acuna, Ling, Kar, Fidler, Efficient Annotation of Segmentation Datasets with Polygon-RNN++, CVPR'18



Total images	160K
Total instances	1.5M
Total hours	10K

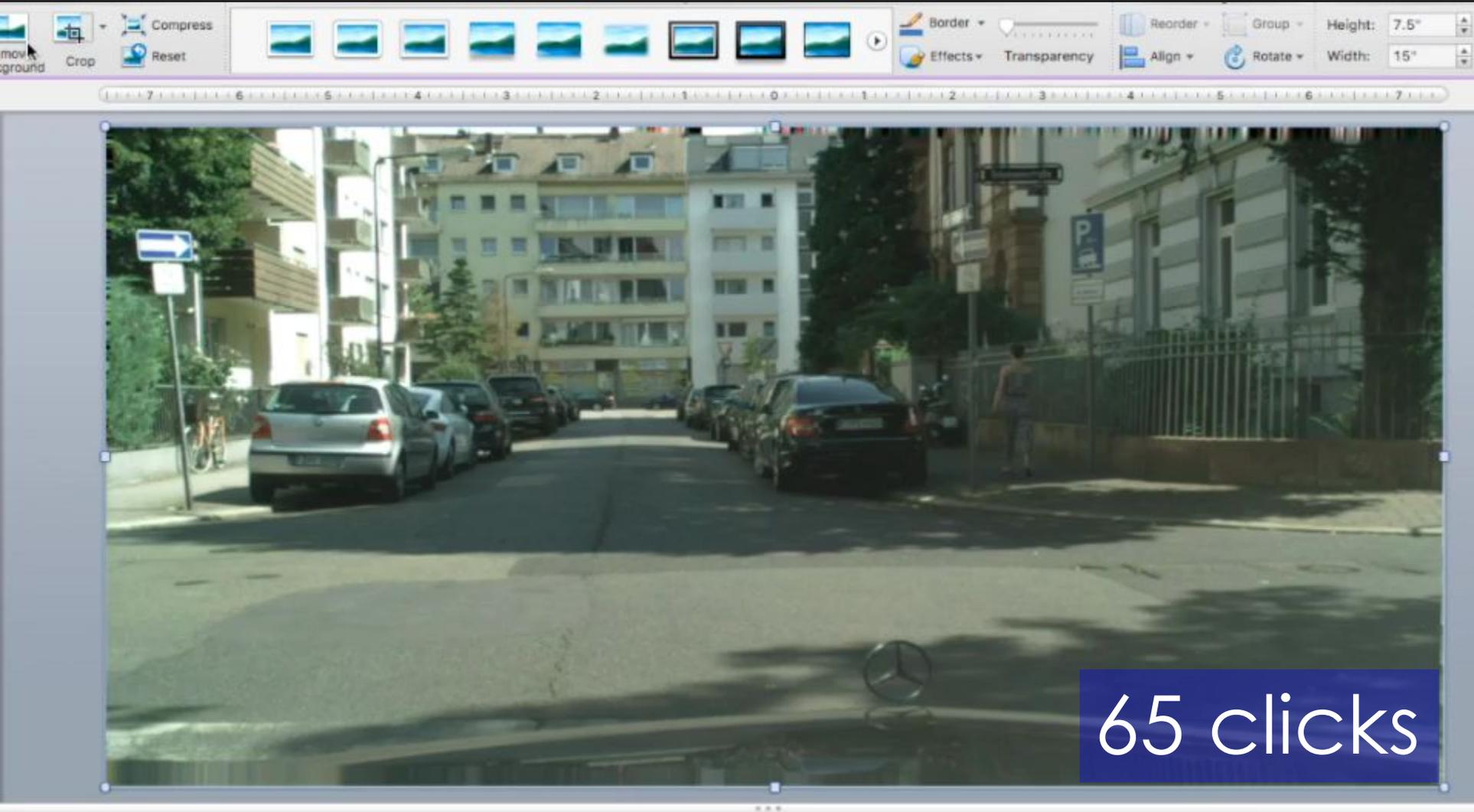


Human-Machine Collaboration



Can we significantly speed up the annotation process?

Interactive Segmentation



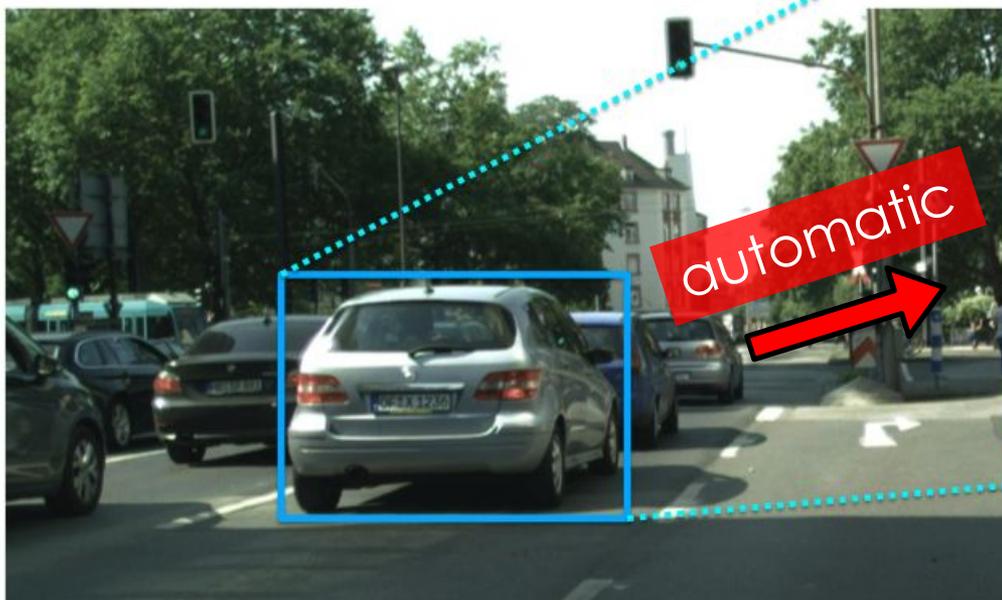
Disclaimer: Recorded with a few year old software

Towards Large-Scale Annotation: Polygon-RNN



Polygon-RNN

Interactive Object Annotation Tool



interactive

 Add box

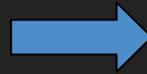
 To correct the prediction, drag and drop a point

Polygon RNN: Model

Input



Output

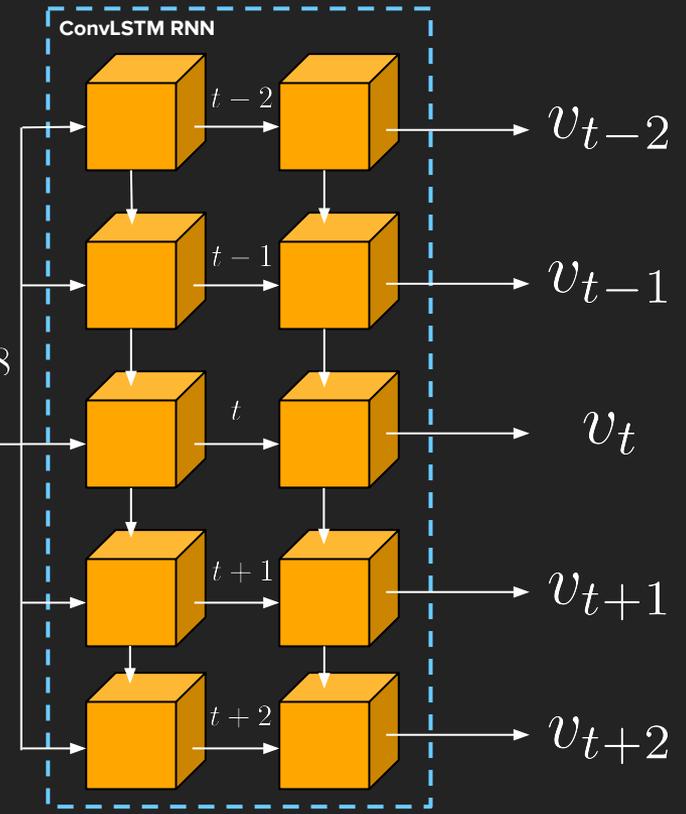
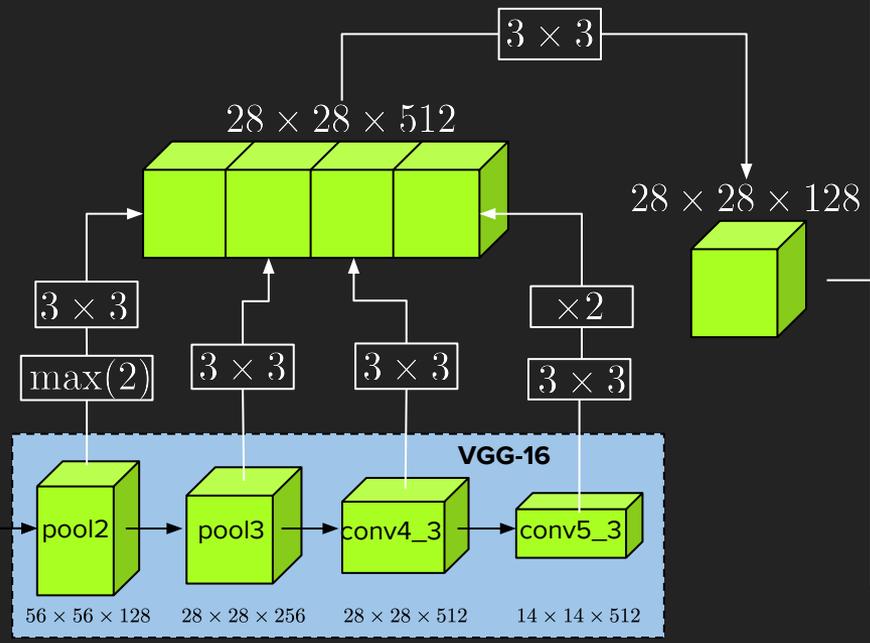


A sequence of vertices
 $\{v_1, \dots, v_n\}, \quad v_i \in \mathbb{R}^2$
outlining the instance

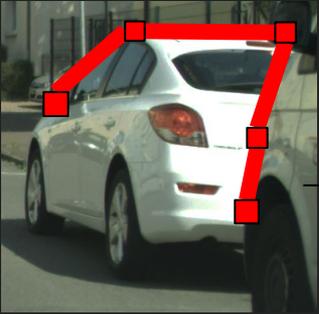
Polygon RNN: Model



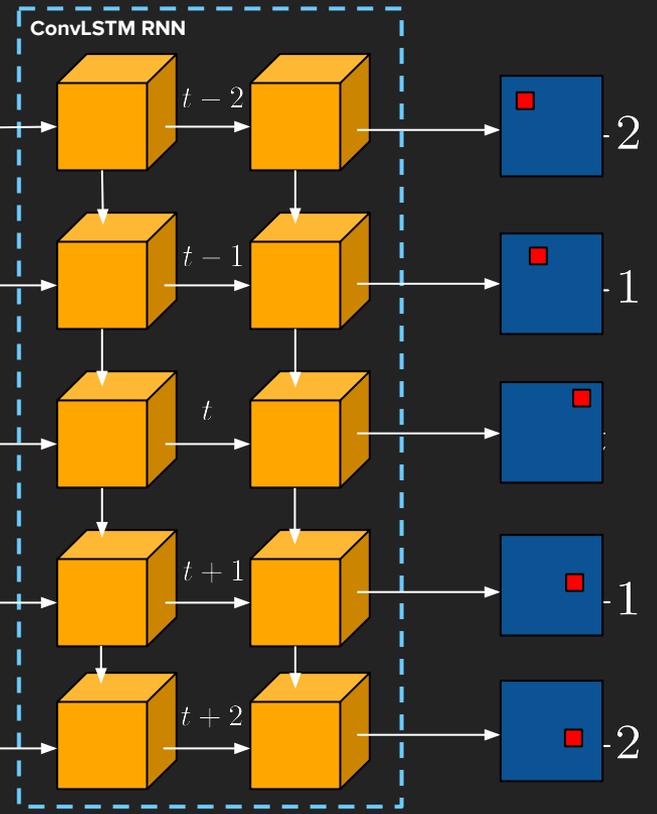
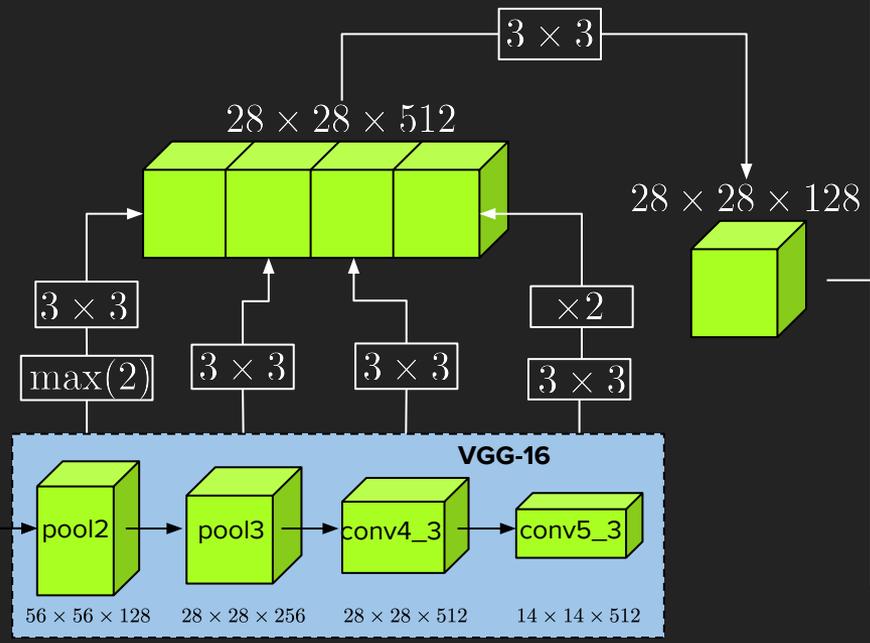
$256 \times 256 \times 3$



Polygon RNN: Model



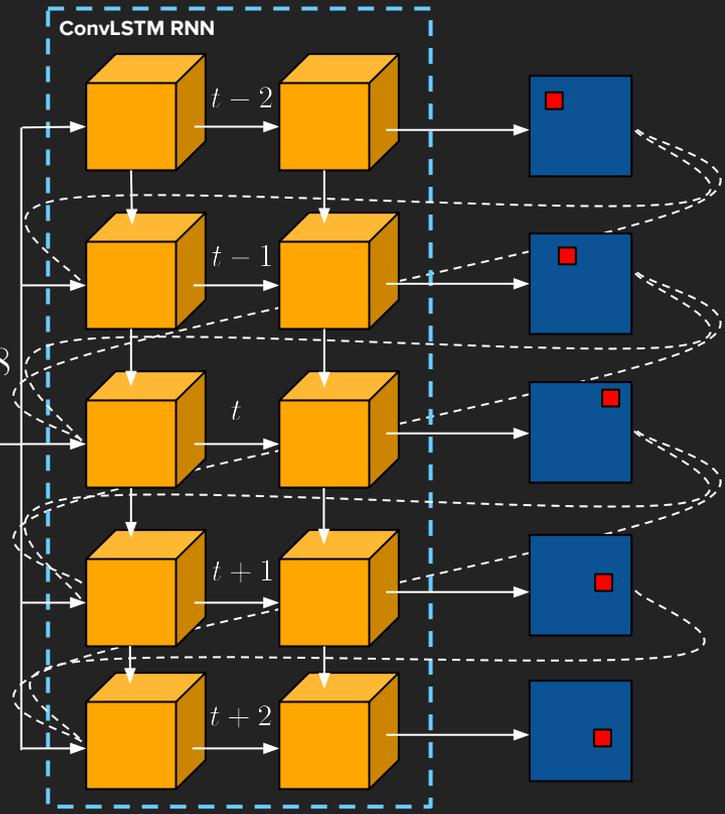
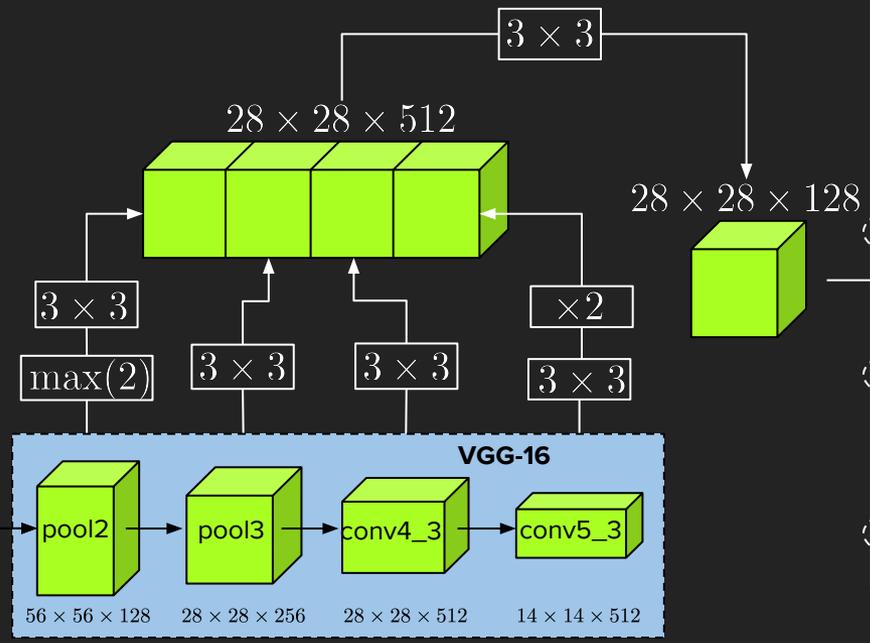
$256 \times 256 \times 3$



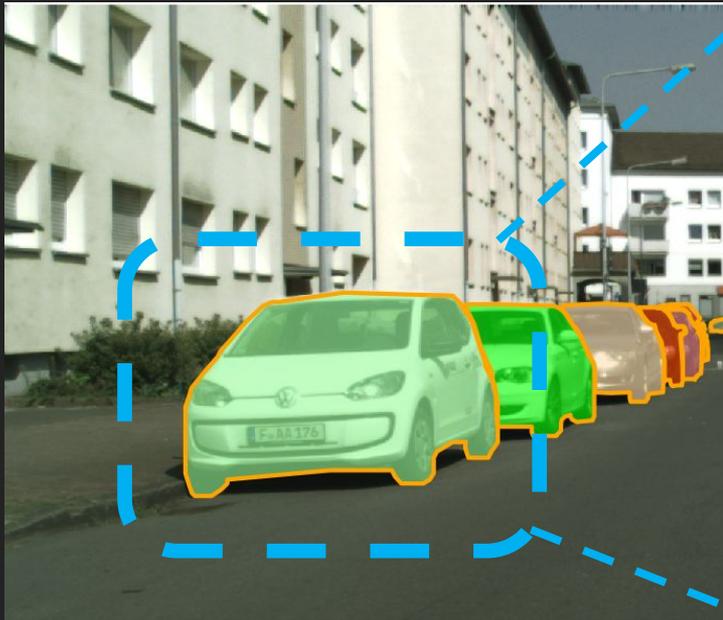
Polygon RNN: Model



$256 \times 256 \times 3$



Upscaling



coarse



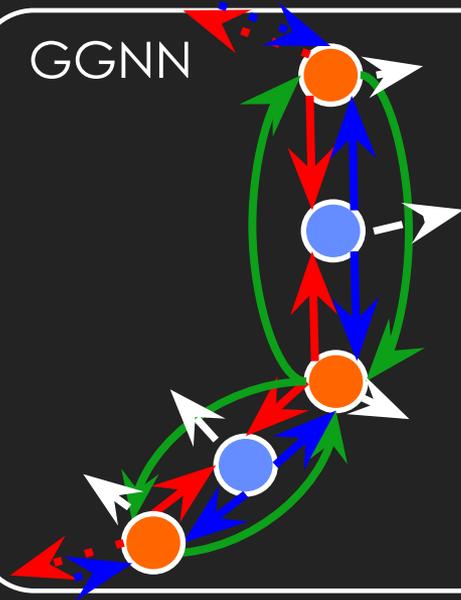
fine

Upscaling

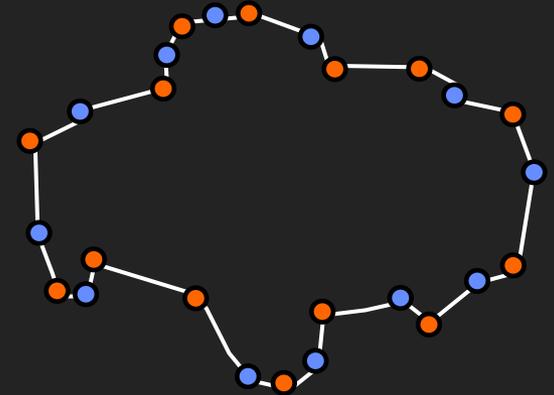
prediction from RNN



GGNN

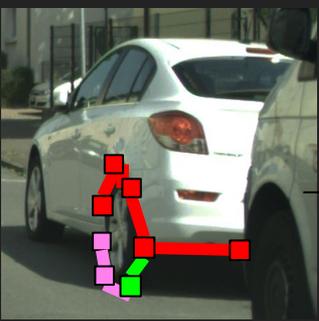


prediction from GGNN

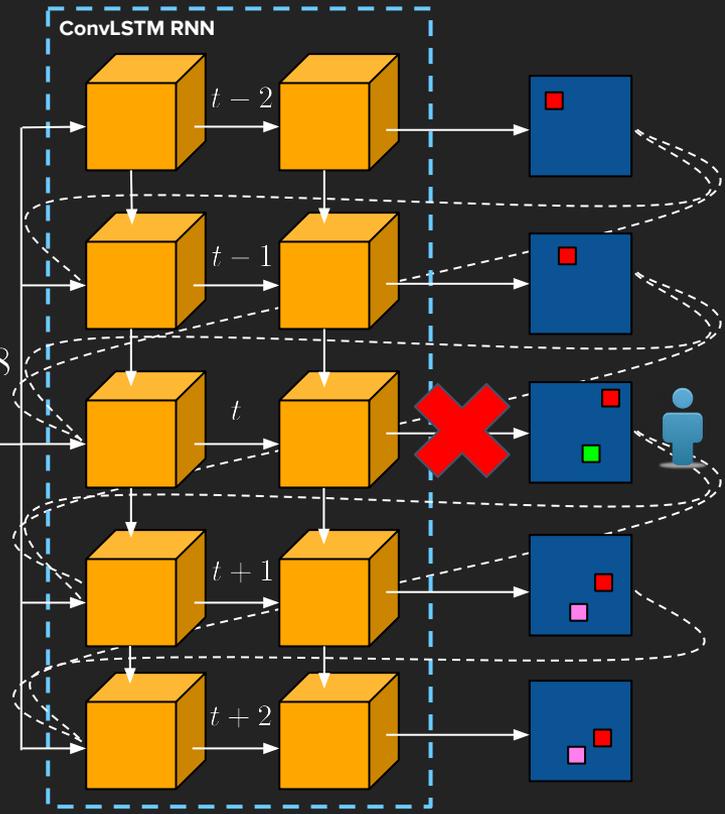
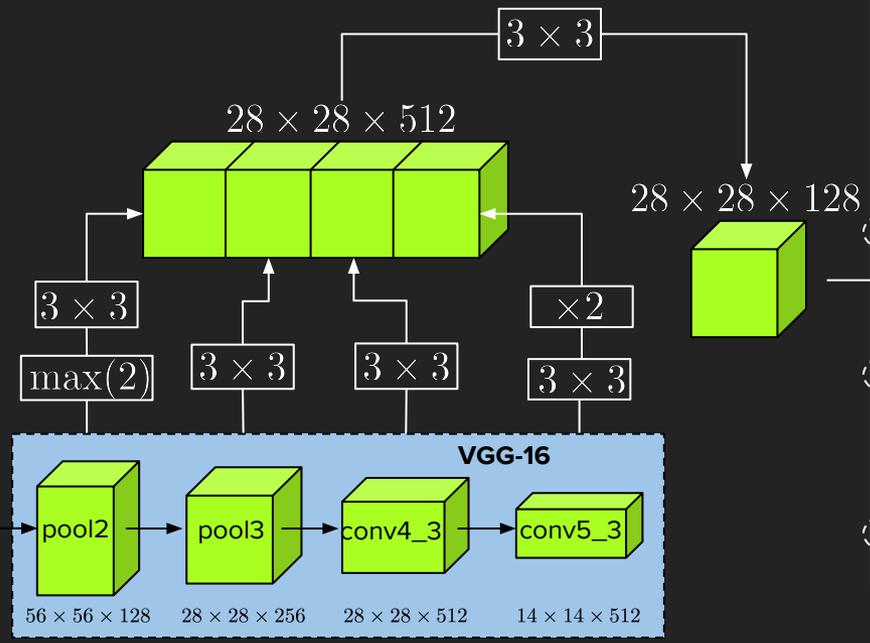


Polygon RNN as the Annotation Tool

Polygon RNN as the Annotation Tool



$256 \times 256 \times 3$



Polygon RNN as the Annotation Tool

Efficient Annotation of Segmentation Datasets with Polygon-RNN++

David Acuna*^{1,3}

Huan Ling*^{1,2}

Amlan Kar*^{1,2}

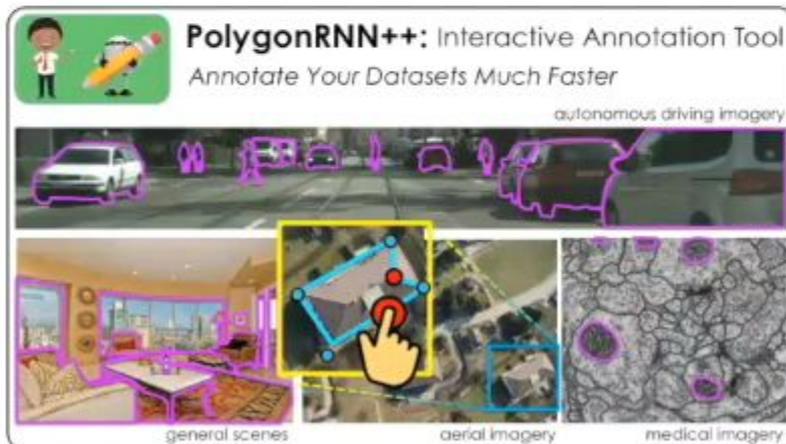
Sanja Fidler*^{1,2}

University of Toronto¹

Vector Institute²

NVIDIA³

{davidj, linghuan, amlan, fidler}@cs.toronto.edu



← equal contribution

research done when D.A. at UofT

Demo



David Acuna

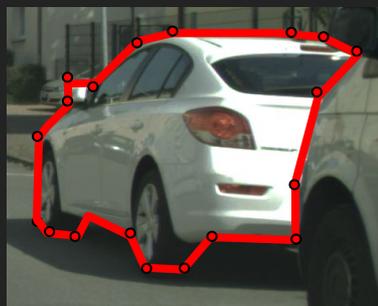
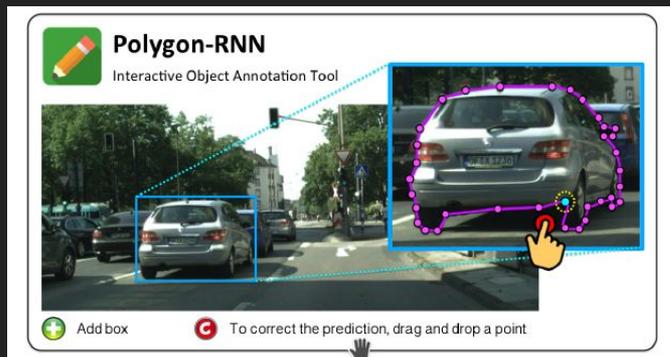


Huan Ling

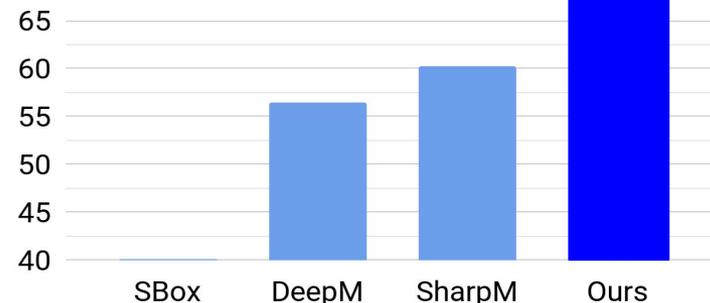


Amlan Kar

0-click Mode



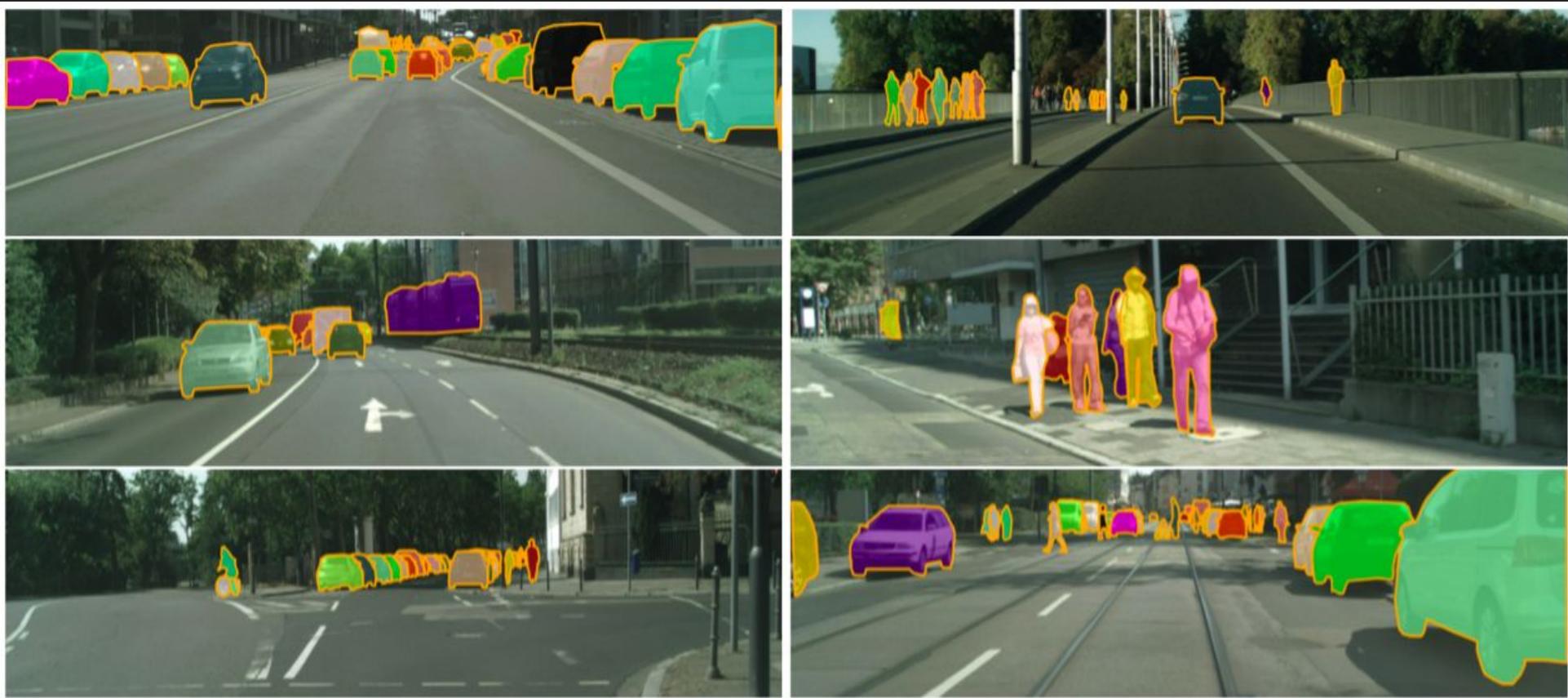
Mean IoU(%) on Cityscapes



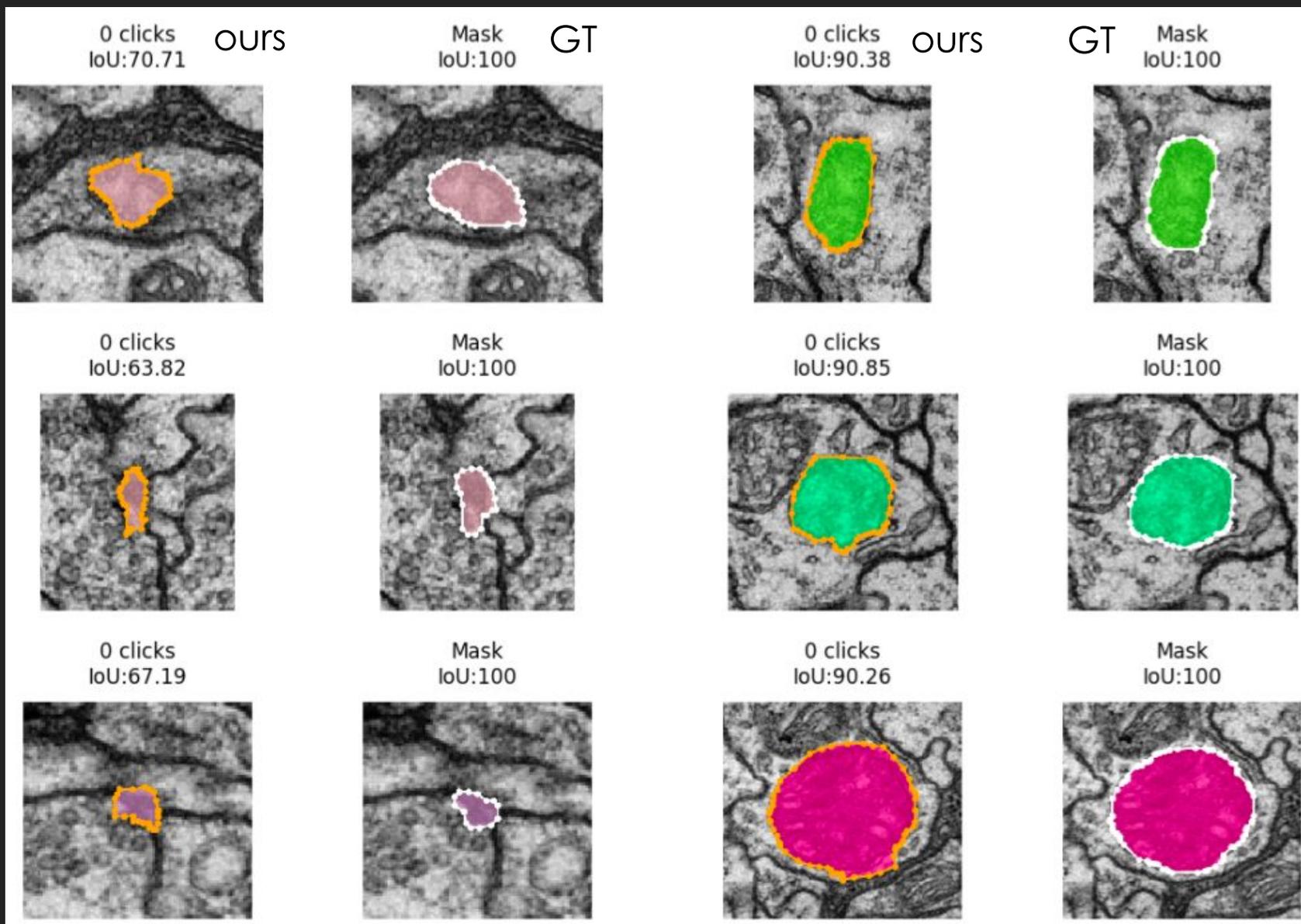
Model	Bicycle	Bus	Person	Train	Truck	Motorcycle	Car	Rider	Mean
Square Box	35.41	53.44	26.36	39.34	54.75	39.47	46.04	26.09	40.11
Dilation10	46.80	48.35	49.37	44.18	35.71	26.97	61.49	38.21	43.89
DeepMask [22]	47.19	69.82	47.93	62.20	63.15	47.47	61.64	52.20	56.45
SharpMask [23]	52.08	73.02	53.63	64.06	65.49	51.92	65.17	56.32	60.21
Polygon-RNN [4]	52.13	69.53	63.94	53.74	68.03	52.07	71.17	60.58	61.40
Residual Polygon-RNN	54.86	69.56	67.05	50.20	66.80	55.37	70.05	63.40	62.16
+ Attention	56.47	73.57	68.15	53.31	74.08	57.34	75.13	65.42	65.43
+ RL	57.38	75.99	68.45	59.65	76.31	58.26	75.68	65.65	67.17
+ Evaluator Network	62.34	79.63	70.80	62.82	77.92	61.69	78.01	68.46	70.21
+ GGNN	63.06	81.38	72.41	64.28	78.90	62.01	79.08	69.95	71.38

0-click Mode

- Annotator draws the box



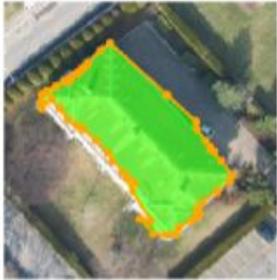
0-click Mode



0-click Mode

0 clicks
IoU:86.65

ours



5 clicks
IoU:100

GT



0 clicks
IoU:88.68

ours



5 clicks
IoU:100

GT



0 clicks
IoU:91.76

5 clicks
IoU:100



0 clicks
IoU:87.55

4 clicks
IoU:100



0 clicks
IoU:88.35

5 clicks
IoU:100

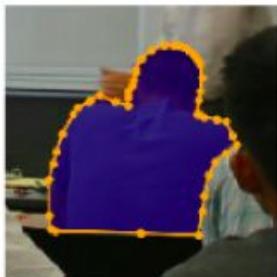


0 clicks
IoU:90.31

5 clicks
IoU:100



0-click Mode

0 clicks
IoU:77.65 ours15 clicks
IoU:100 GT0 clicks
IoU:87.20 ours34 clicks
IoU:100 GT0 clicks
IoU:87.4414 clicks
IoU:1000 clicks
IoU:82.0026 clicks
IoU:1000 clicks
IoU:90.1525 clicks
IoU:1000 clicks
IoU:90.829 clicks
IoU:100

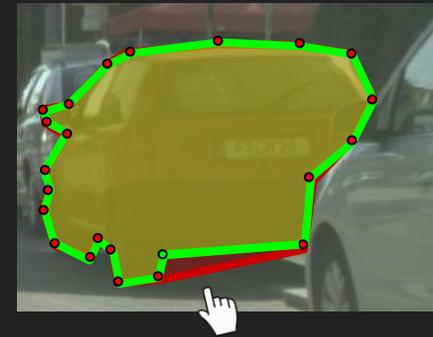
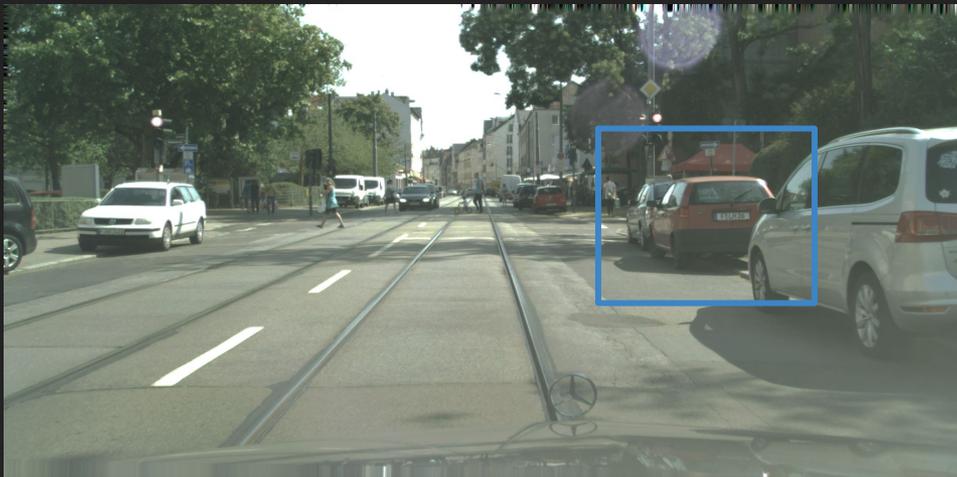
Out-of-Domain (0-click) Prediction

- Annotator draws the box

Model	ADE	Rooftop	Cardiac MR	ssTEM
DeepMask[22]	59.74	15.82	60.70	31.21
SharpMask[23]	61.66	18.53	69.33	46.67
Ours w/o GGNN	70.21	65.03	80.55	53.77
Ours w/ GGNN	71.82	65.67	80.63	53.12

Annotation Mode

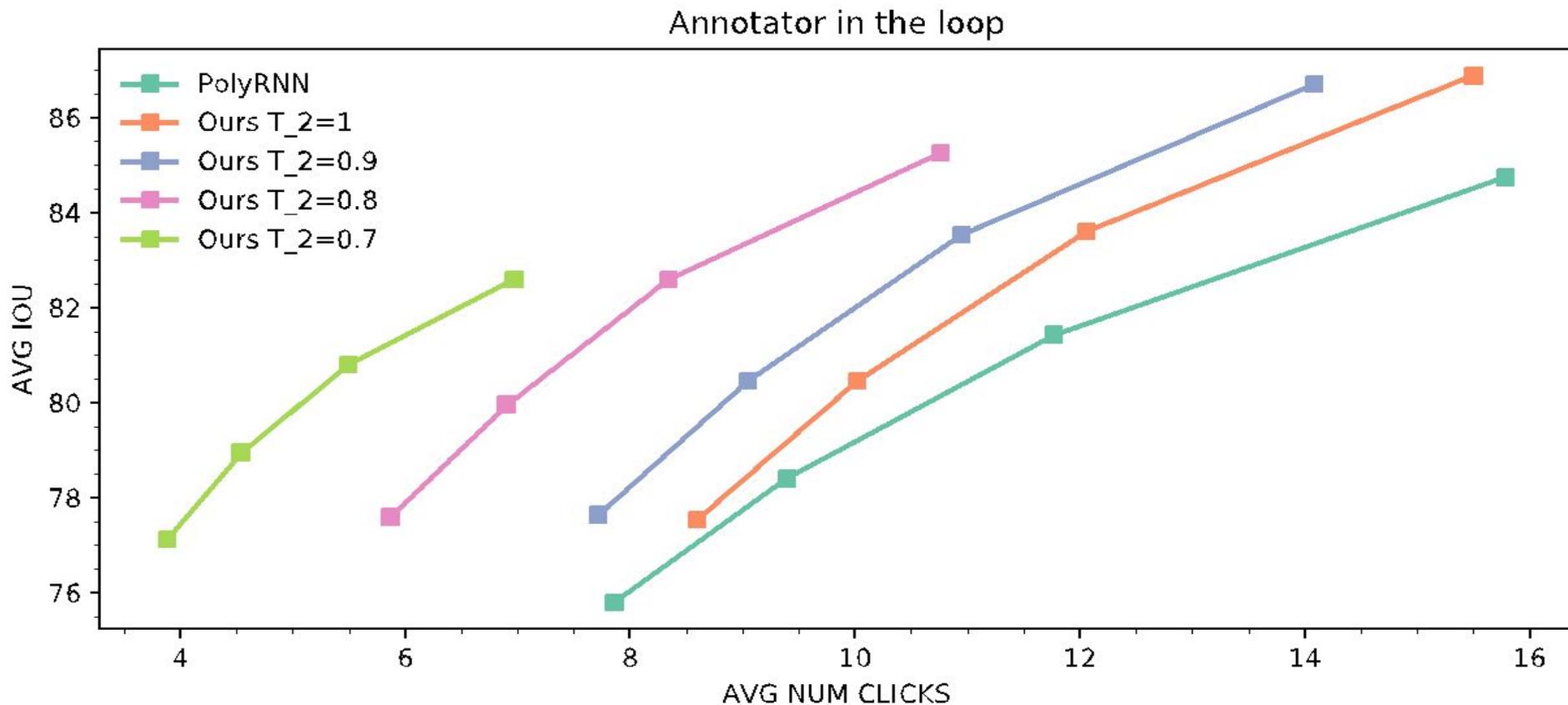
We **simulate** a human annotator correcting vertices



Automatic correction if
distance to GT vertex $>$
threshold T

Evaluation: i) Number of clicks per instance, ii) resulting IoU

Polygon RNN as the Annotation Tool



Human Experiment

Cityscapes

ADE

	Time (s)	IOU (%)	Time (s)	IOU (%)
manual	39.7	76.2	29.2	80.6
Polygon-R NN++	14.7	75.4	19.3	75.9

Grabcut

42.2

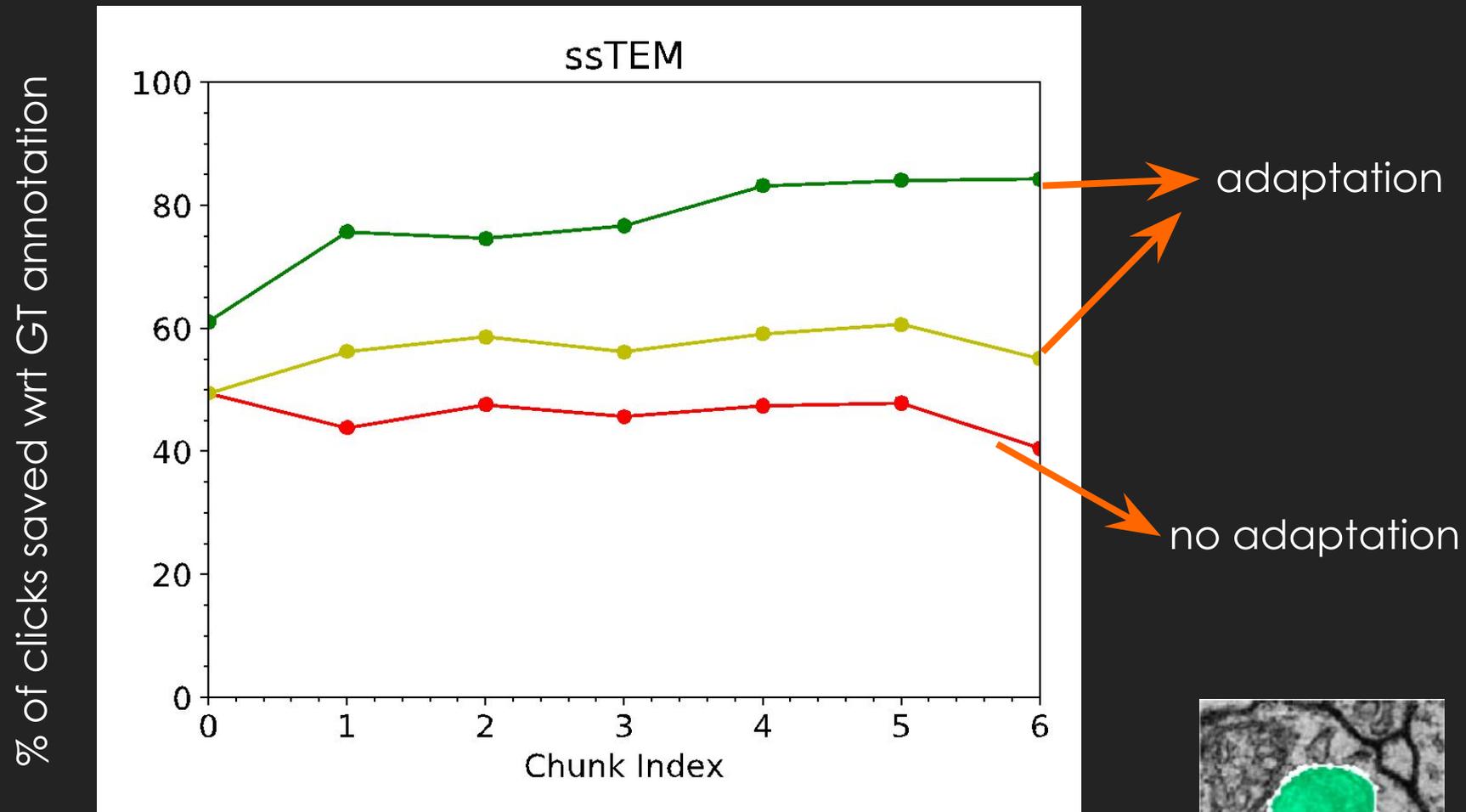
70.7

(model trained on Cityscapes,
no finetuning)

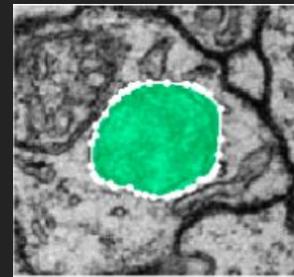
Adapting to New Domains

```
for i = 1 ... #chunks
    PredictAndCorrect
    fine-tune on (all new domain) data
end
```

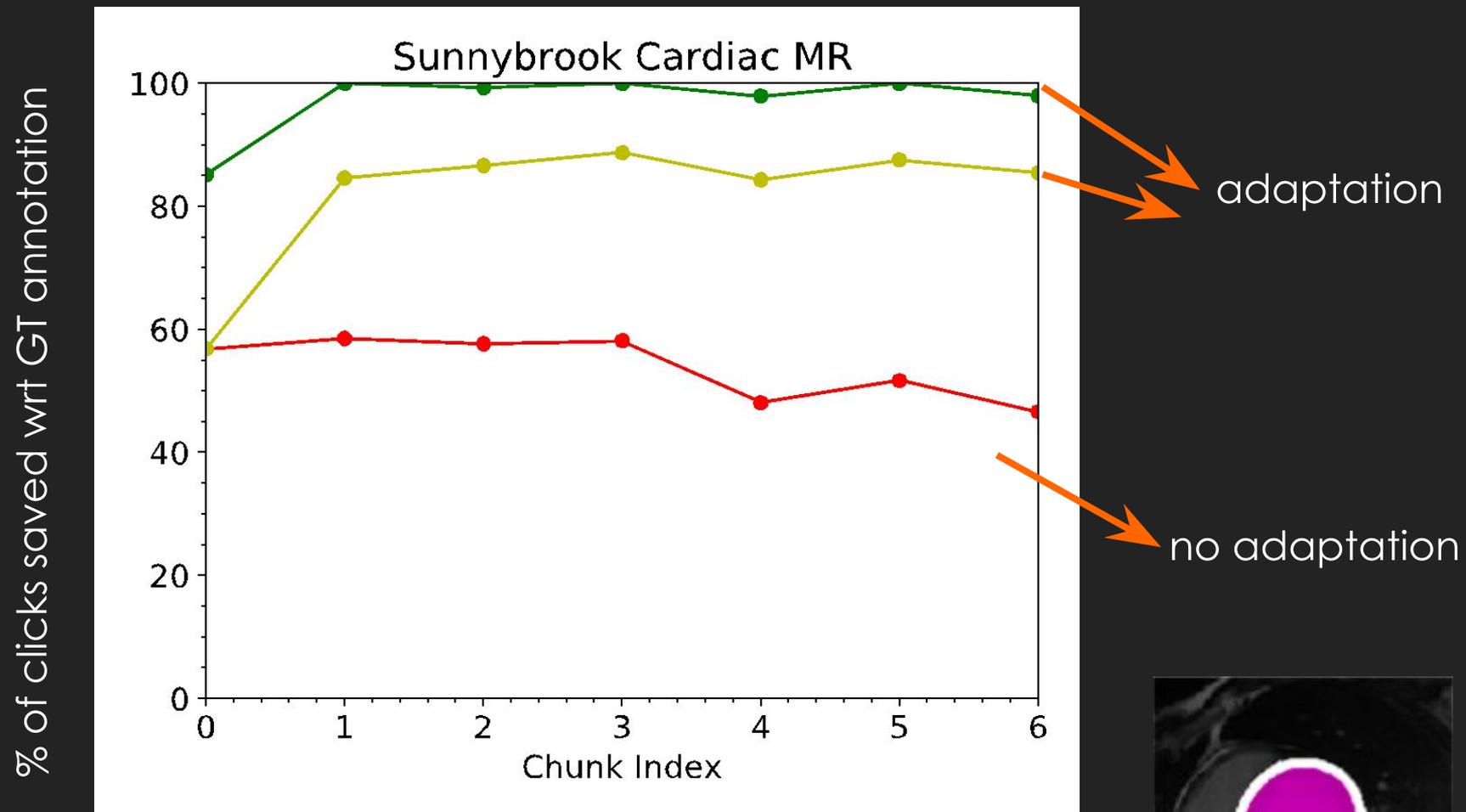
Adapting to New Domains



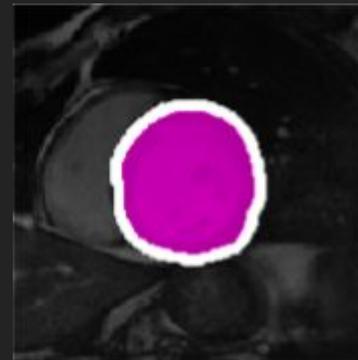
medical dataset



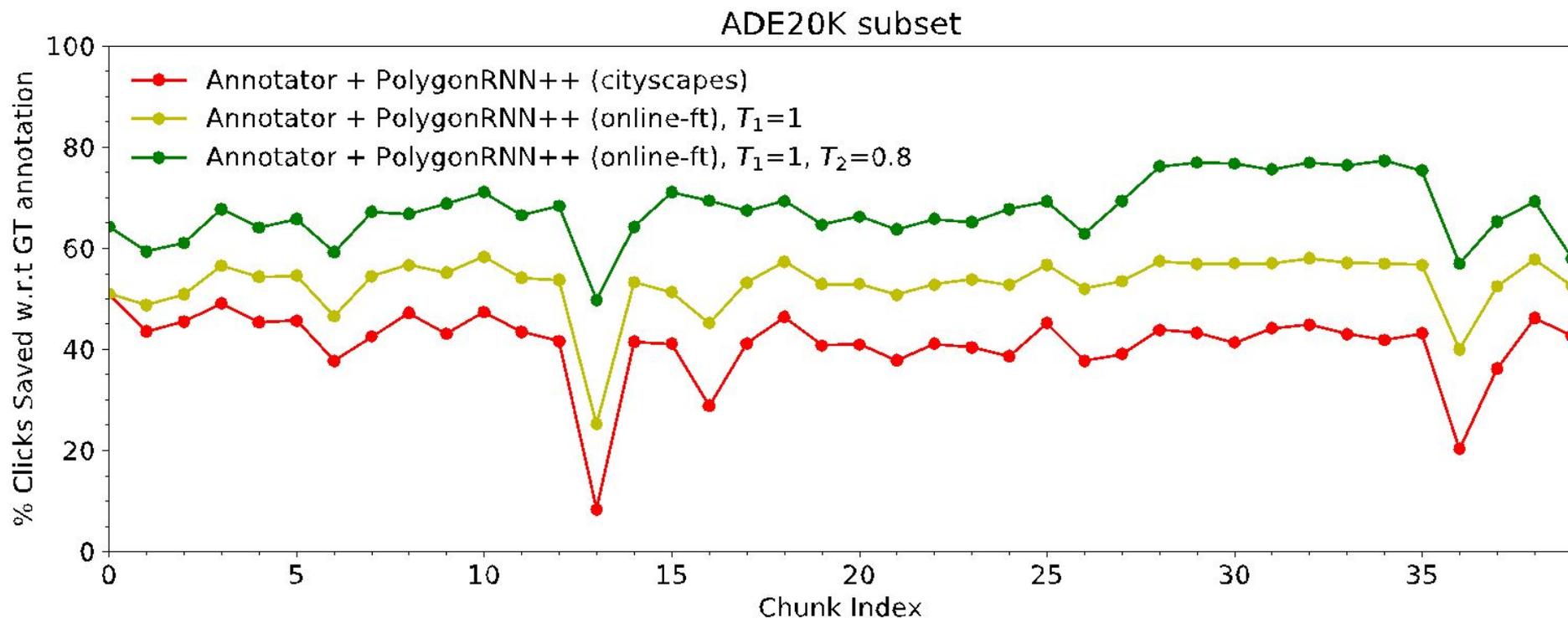
Adapting to New Domains



medical dataset



Adapting to New Domains



generic imagery



Training on Mapping Data



Acuna, Ling, Kar, Fidler, CVPR'18

Summary and Future work

- Video annotation
- 3D annotation

Polygon-rnn website:

<http://www.cs.toronto.edu/polyrnn>

Thank You