Deep Learning on Graphs ——CNN, RNN and GNN

Huan Ling

University of Toronto & Vector Institute & Nvidia Research Lab

Toronto 9th/Oct

VECTOR INSTITUTE | INSTITUT VECTEUR

UNIVERSITY OF TORONTO

About

- The Vector Institute is an independent, not-for-profit entity dedicated to research in artificial intelligence with a focus on excellence in machine learning and deep learning.
- Vector launched in 2017 with support from the Government of Ontario, Government of Canada, and the private sector.

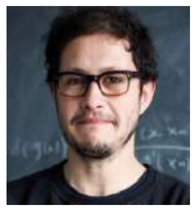
Researchers


Geoffrey Hinton
Chief Scientific Advisor


Richard Zemel
Research Director


Alán Aspuru-Guzik


Jimmy Ba


Juan Carrasquilla


David Duvenaud


Murat Erdogdu


Amir-massoud Farahmand


Sanja Fidler


David Fleet


Brendan Frey


Marzyeh Ghassemi


Anna Goldenberg


Roger Grosse


Alireza Makhzani

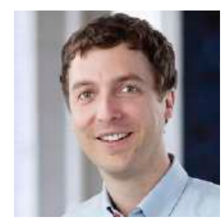

Quaid Morris


Sageev Oore


Pascal Poupart


Daniel Roy


Frank Rudzicz


Graham Taylor


Raquel Urtasun

Over the next few years it is expected that Vector will grow to include several hundred researchers, students, engineers and others.

Today

Deep Learning & Computer Vision on Graphs

Citation Network , Social Network

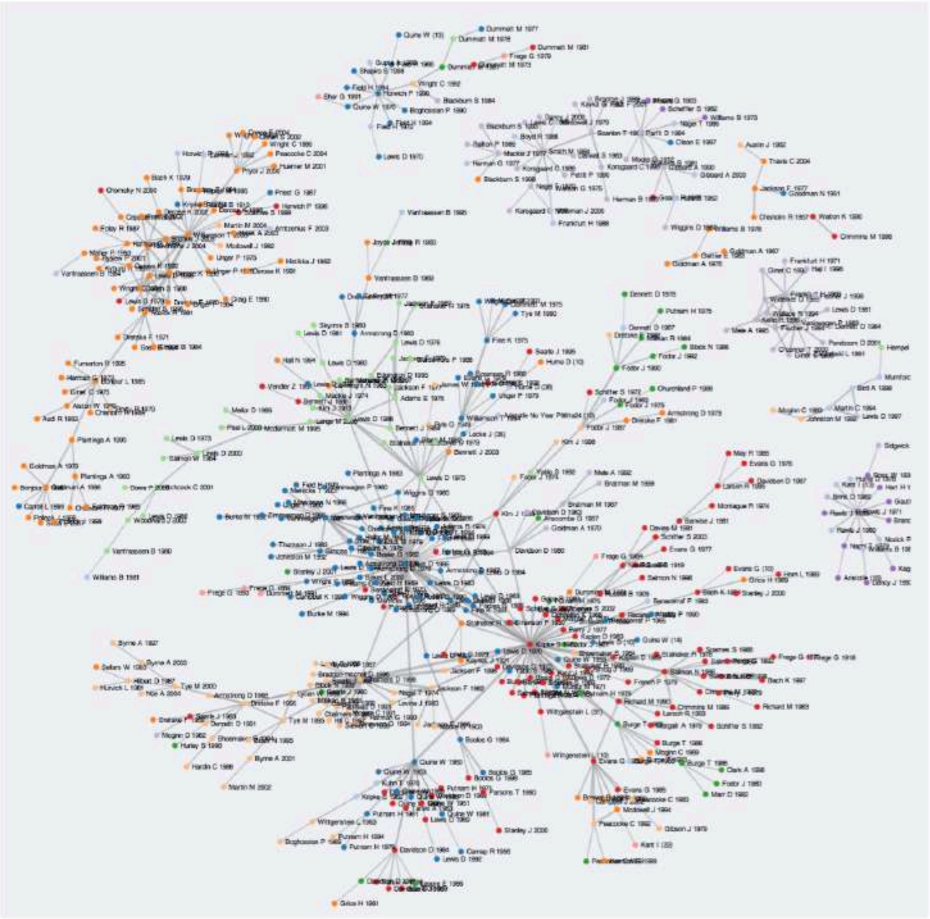


Figure: Citation network for Philosophy [1]

Image credit from Kieran Healy

Reinforcement Learning

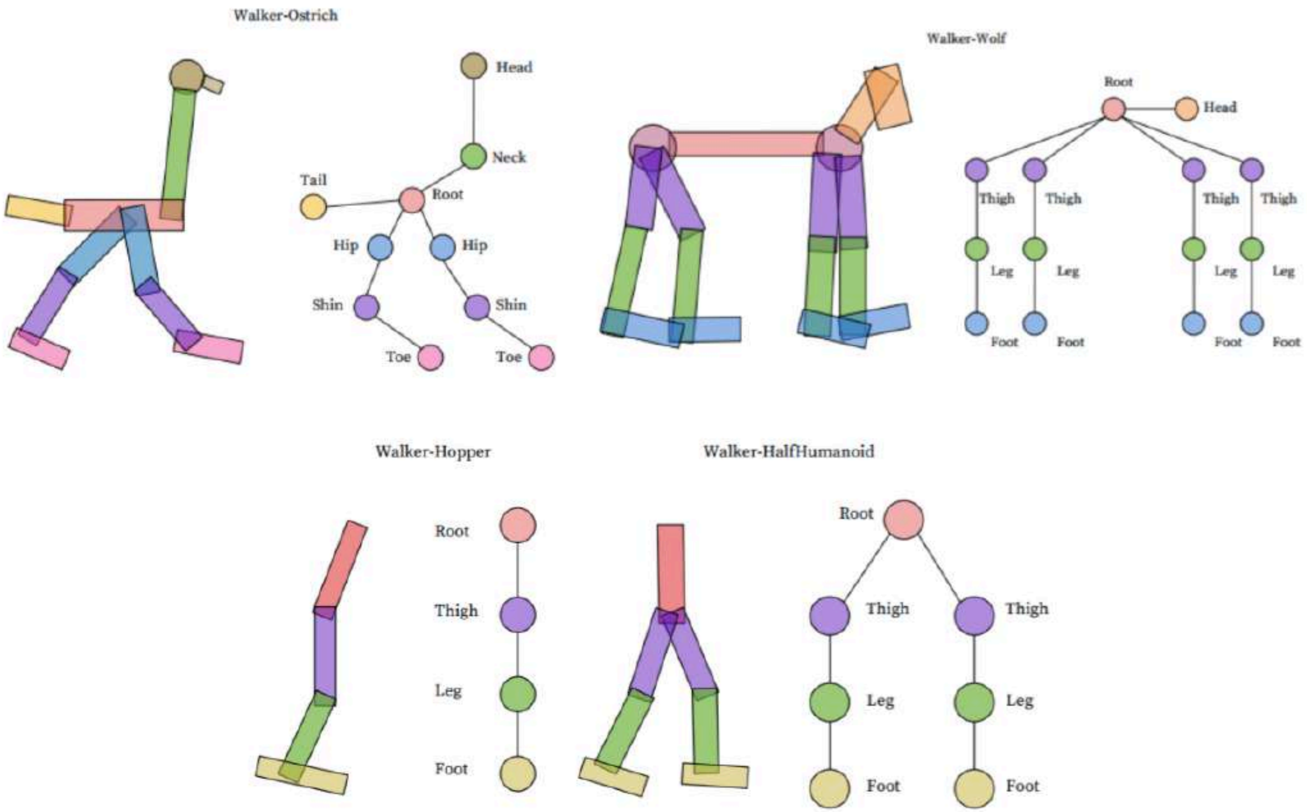
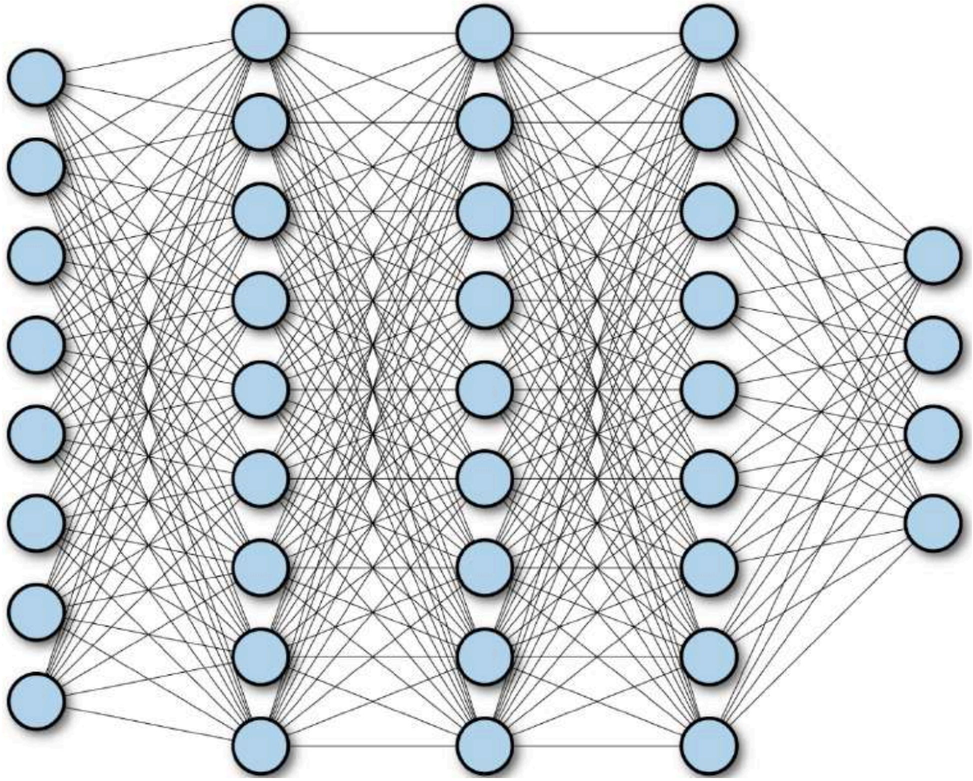


Figure: NerveNet [3]

Skeleton as a graph

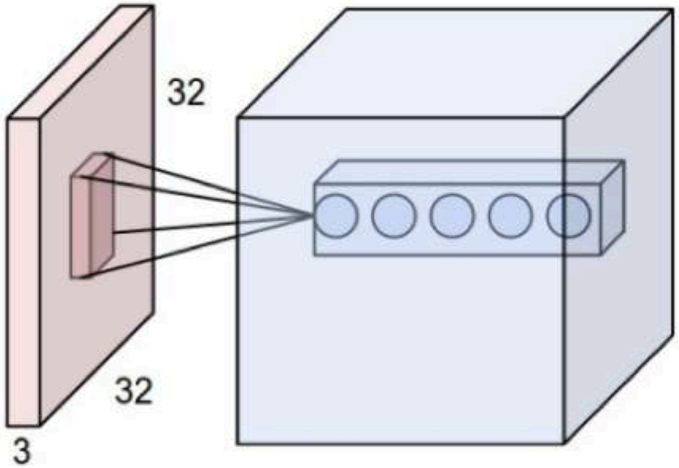
Image credit from (Wang et al., 2017)

Quick Review: Neural Networks



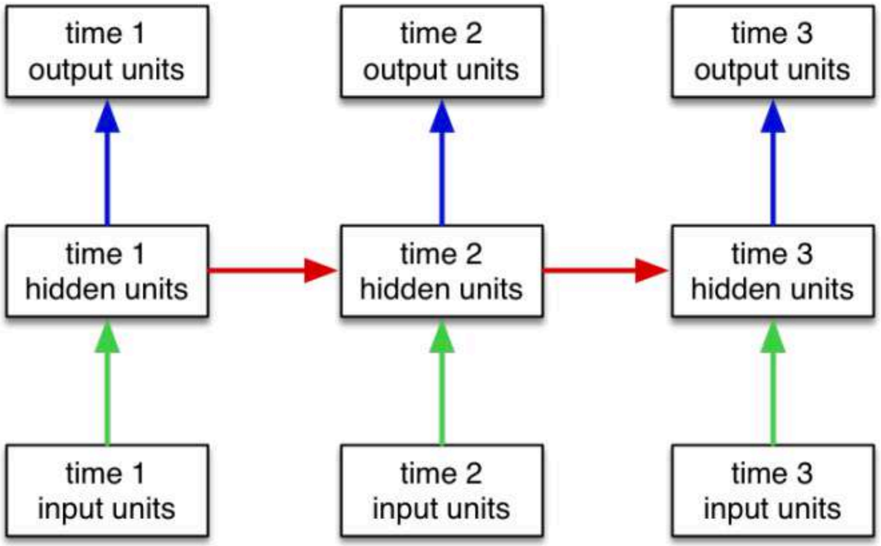
Fully connected neural networks

Quick Review: Neural Networks



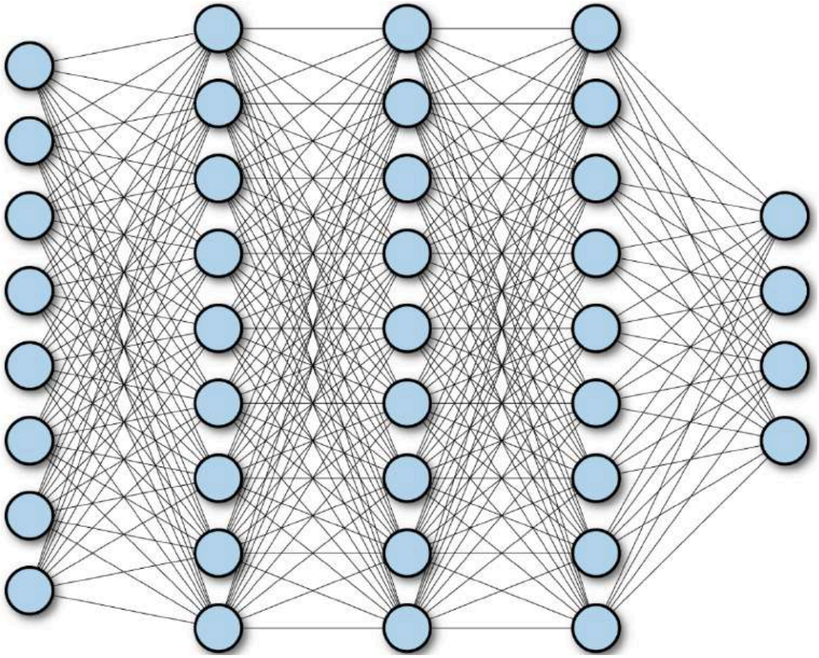
Convolutional neural network

Quick Review: Neural Networks



Recurrent neural nets

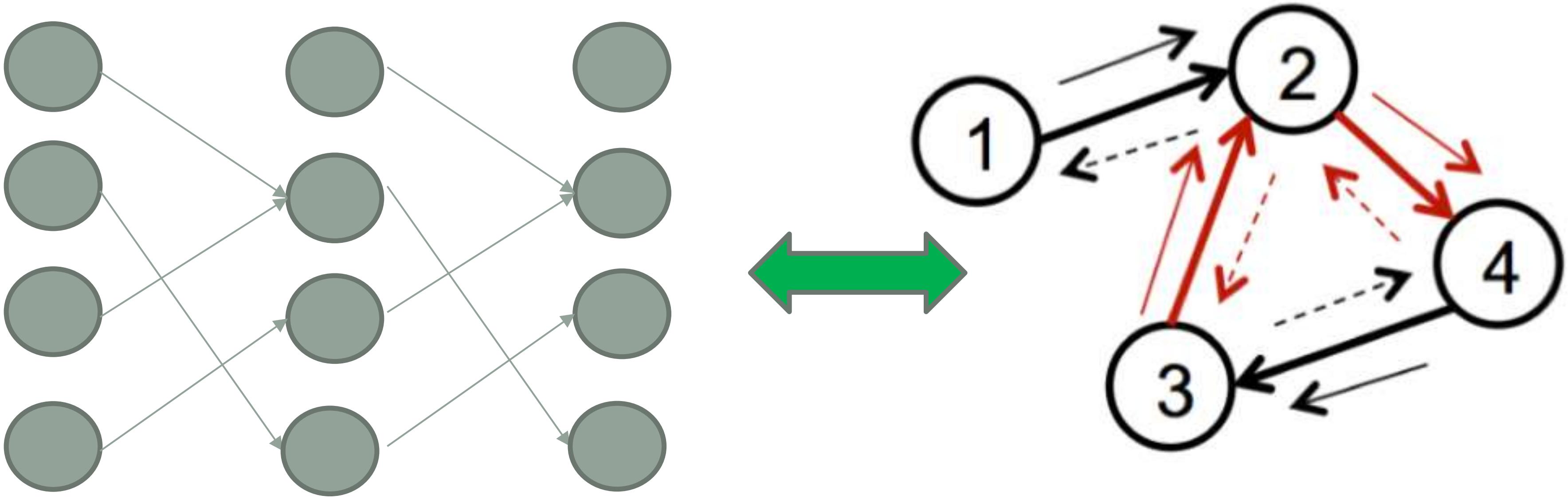
Quick Review: Neural Networks



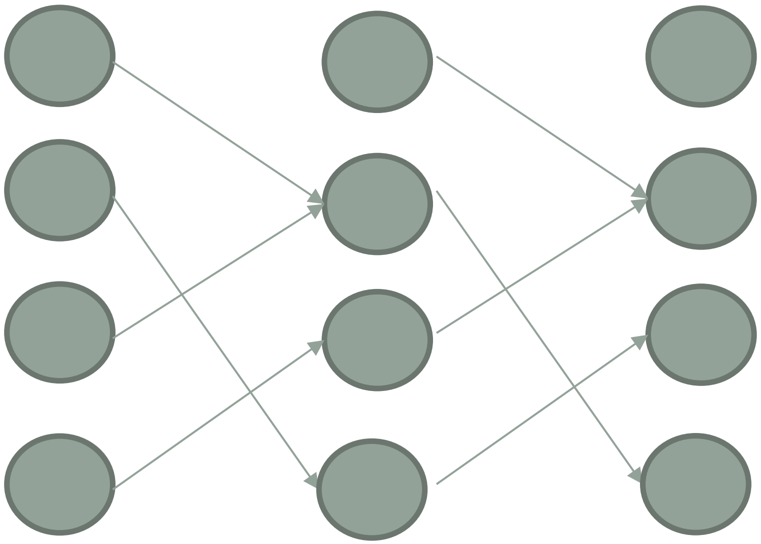
What if each node at input layer has it owns semantic meaning?
(i.e friendship in social network)
What if we want our own link rules?

Neural Network on Graph

Define propogation by pre-defined rules



Graph Convolutional neural network



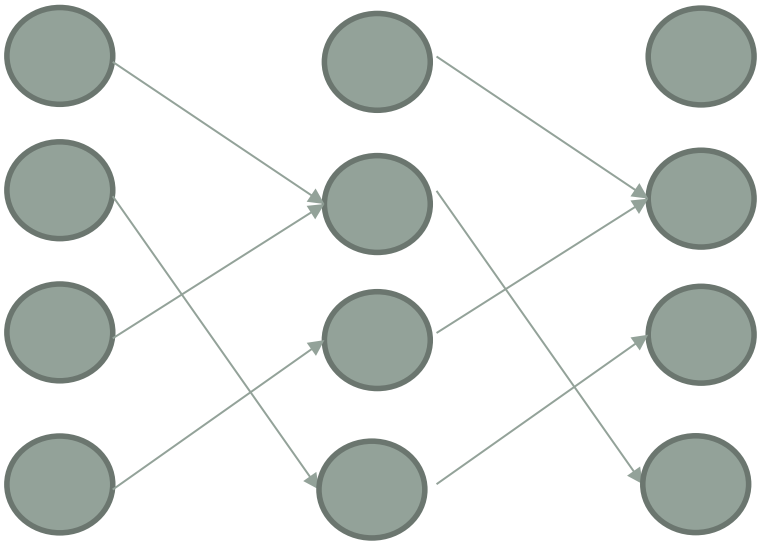
Aggregate node as CNN

N(p) is neighbourhood nodes for p

l is time step

$$f_p^{l+1} = w_0 f_p^l + \sum_{q \in \mathcal{N}(p)} w_1 f_q^l$$

Gated Graph Sequence Neural Networks



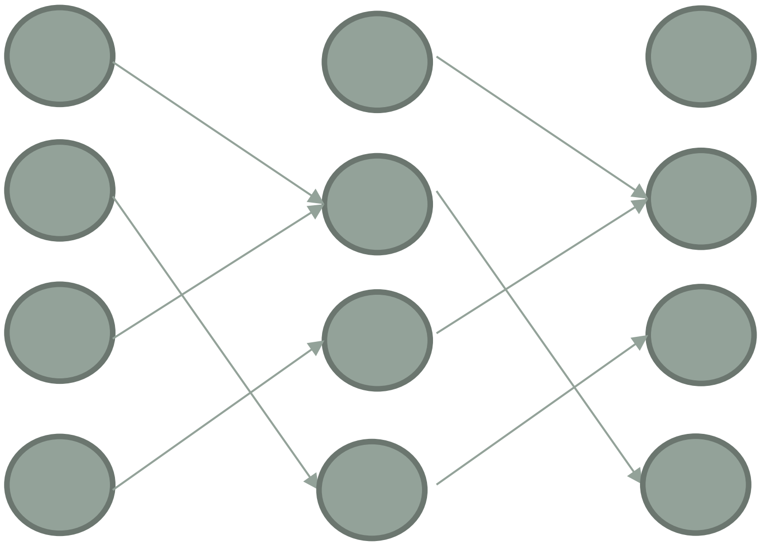
Aggregate node as RNN

N(p) is neighbourhood nodes for p

l *is time step*

$$f_p^{l+1} = w_0 f_p^l + \sum_{q \in N(p)} RNN(q)$$

Graph Neural Networks



Notes:

1. GCN is more suitable for computer vision problem
2. GGNN has the sense of distance and time steps

All details are omited.
Read reference for further details.

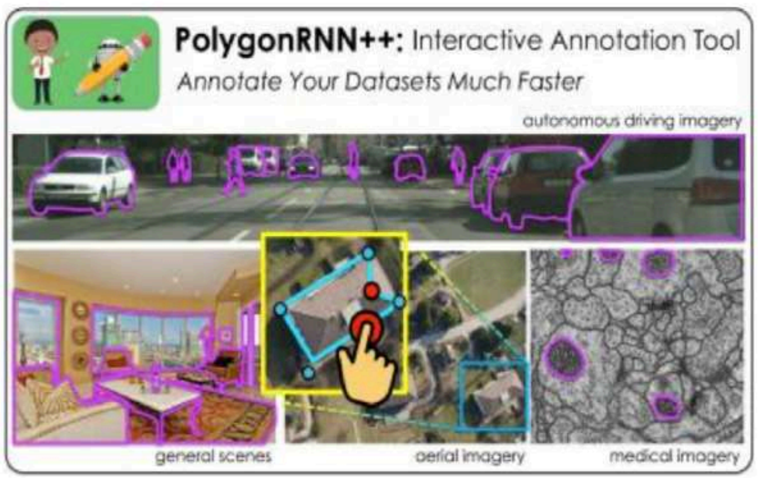
Polygon Image Segmentation

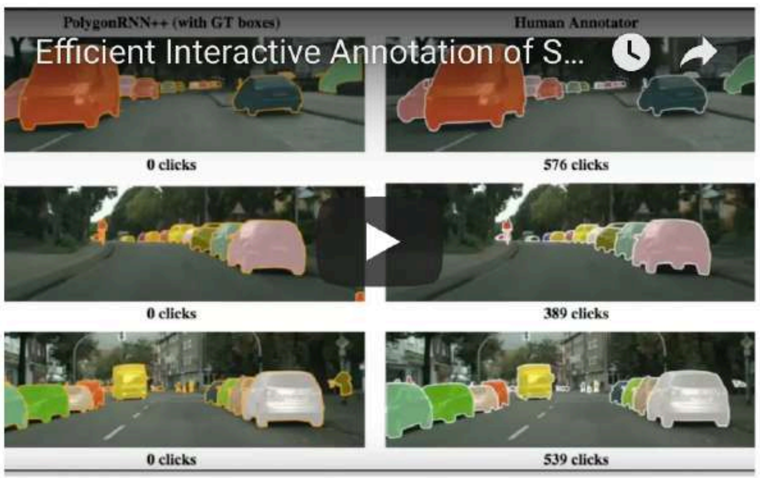
Efficient Annotation of Segmentation Datasets with Polygon-RNN++

David Acuna* [1,3] Huan Ling* [1,2] Amlan Kar* [1,2] Sanja Fidler [1,2]

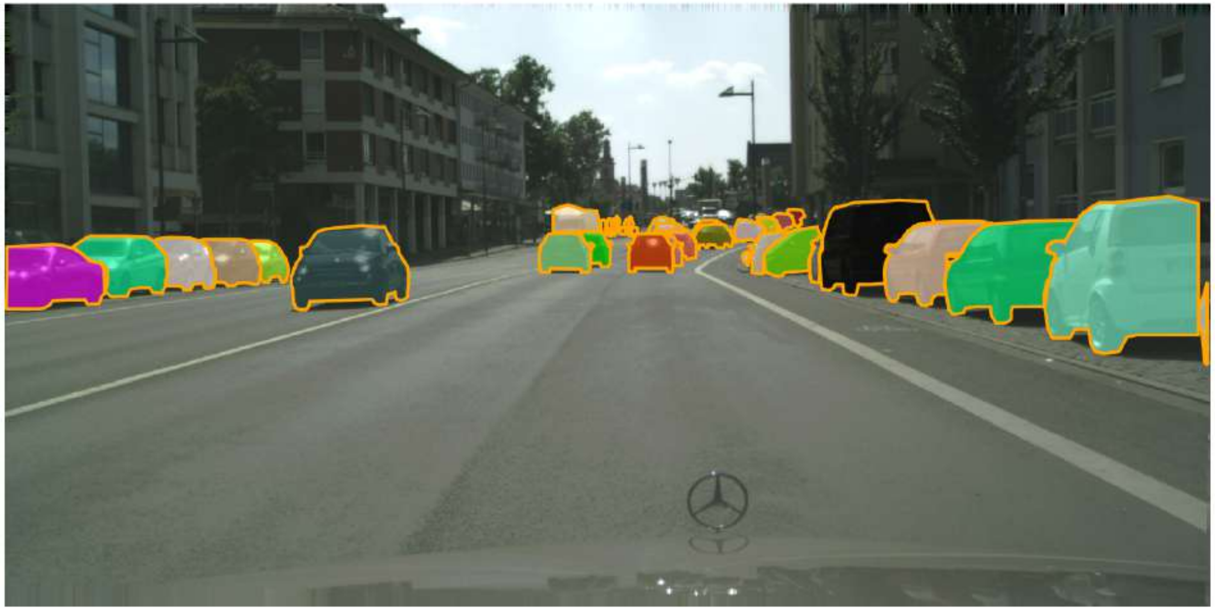
[1] University of Toronto [2] Vector Institute [3] NVIDIA#

CVPR, 2018





Polygon RNN

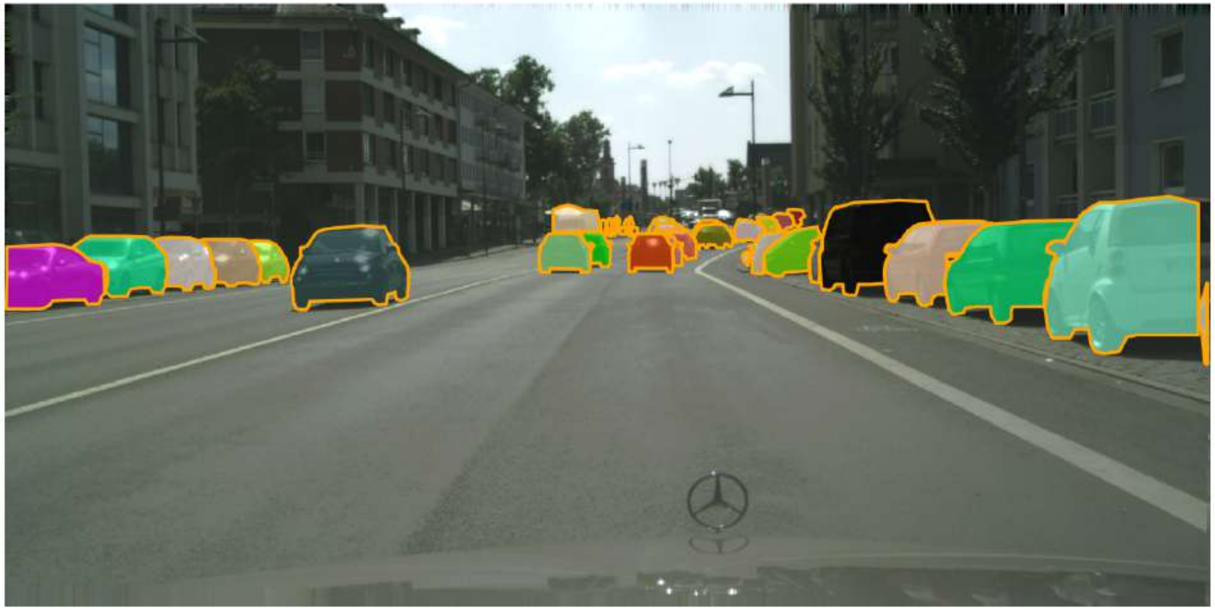


Why Polygon?

1. Human in the loop

2. Deformable

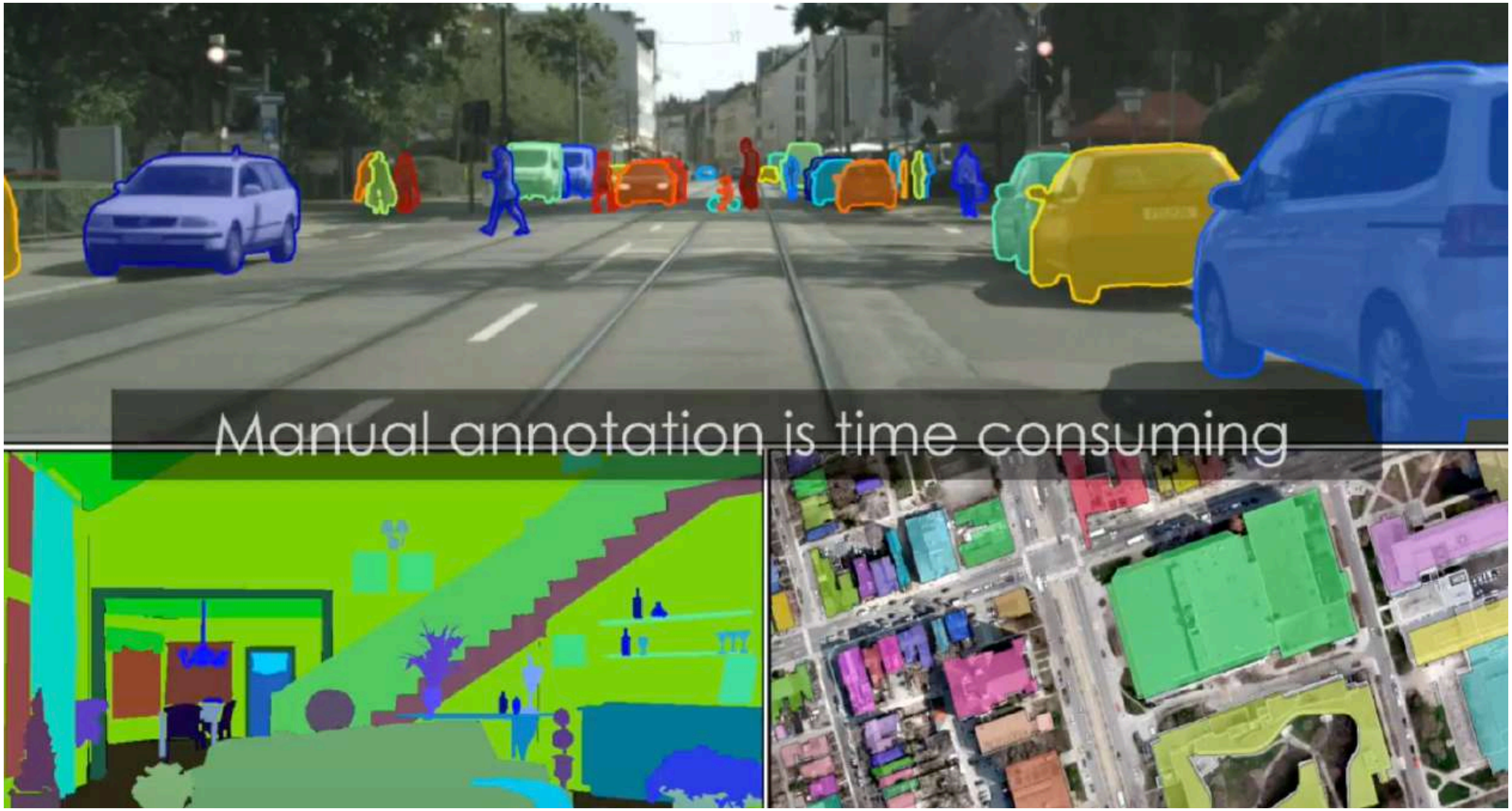
Polygon RNN



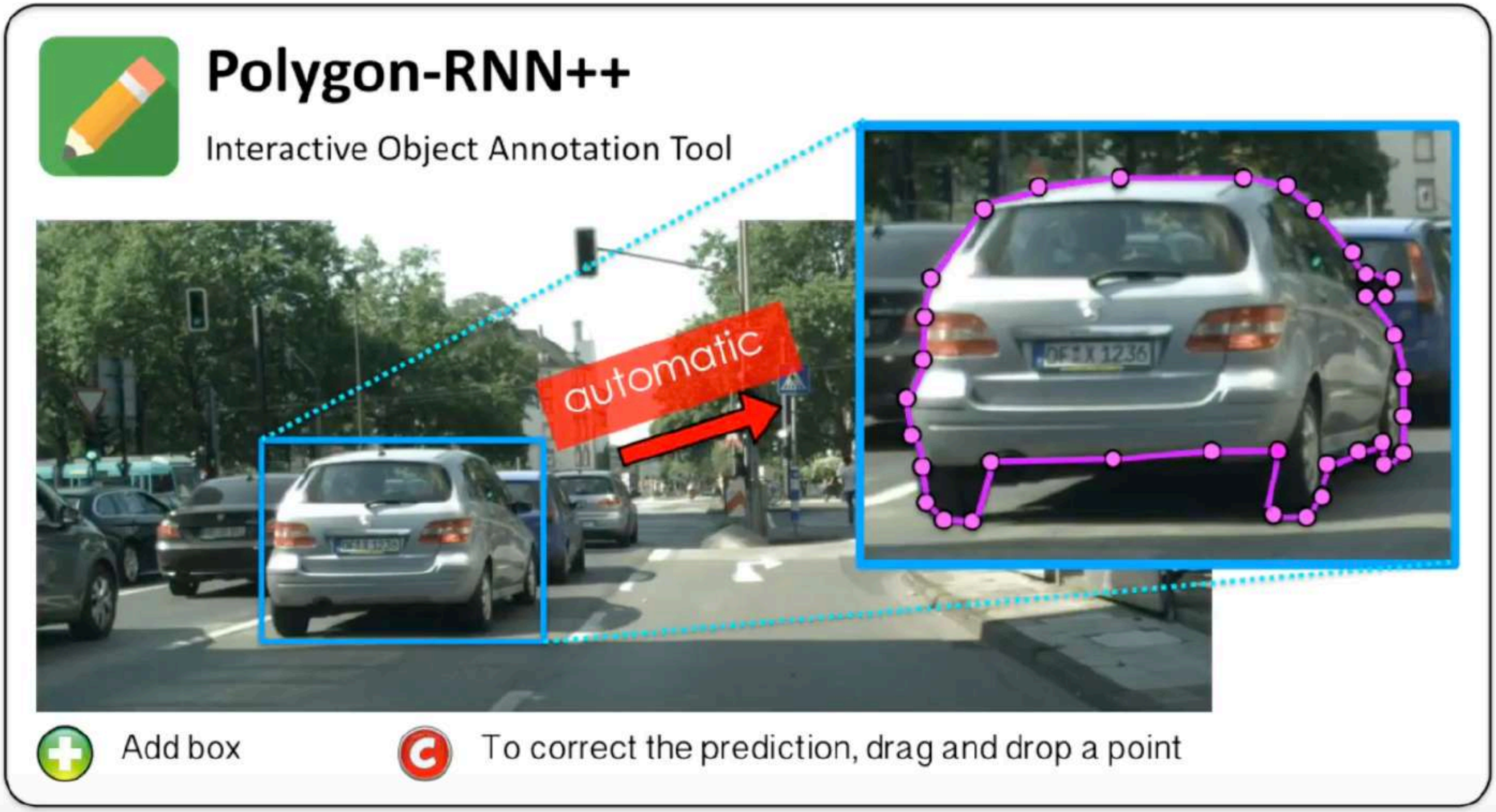
Why Polygon?

1. **Human in the loop**
2. Deformable

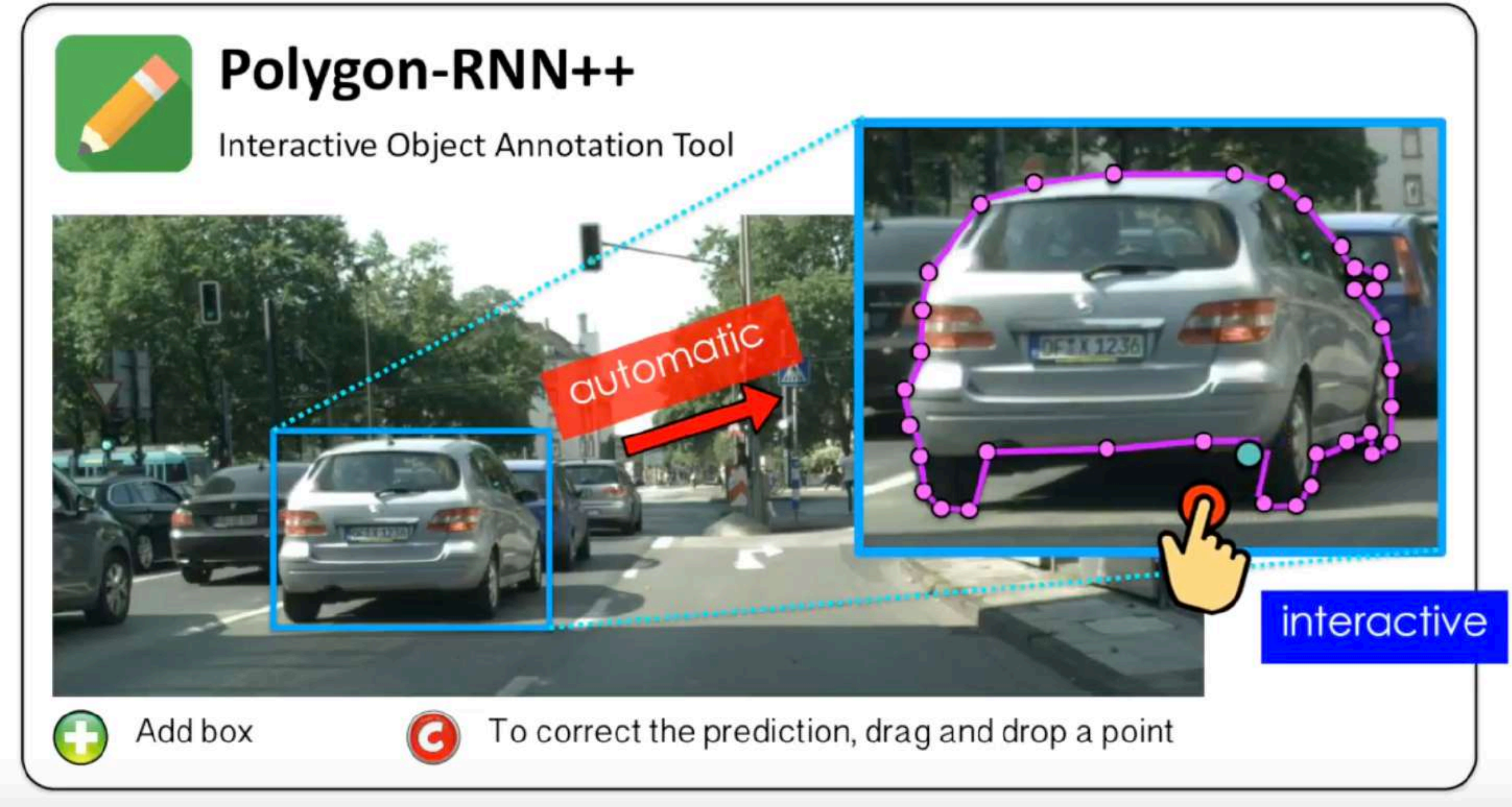
Polygon RNN : Bring human to the loop



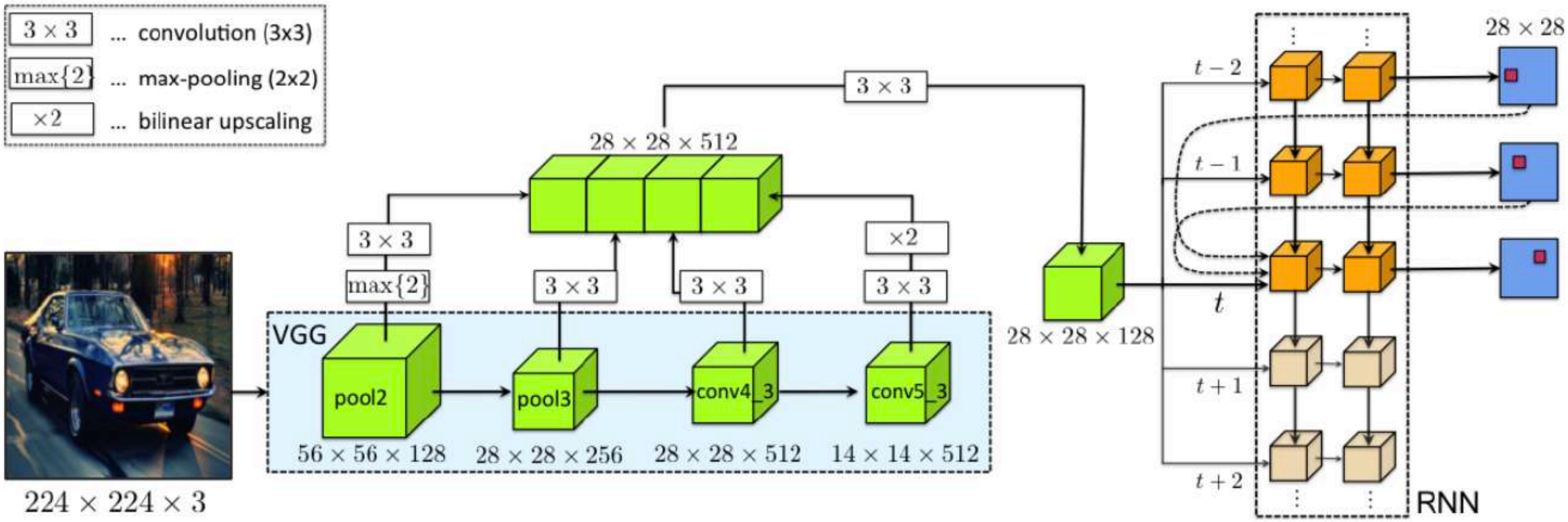
Polygon RNN : Bring human to the loop



Polygon RNN : Bring human to the loop



Polygon RNN : Bring human to the loop



Demo

PolygonRNN++

Interactive Object Annotation

our AI assisted tool can help you annotate objects in images via polygons

Go ahead. Draw a box!

Enter Image URL

Submit URL | Upload Local

New Object (n)

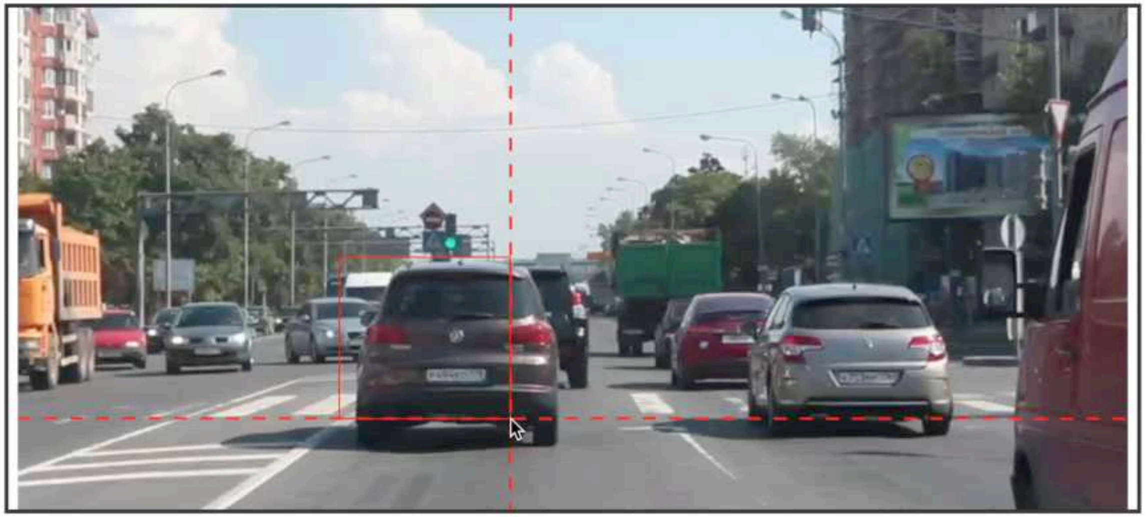
Finish Object (f)

Discard Object (d/del)

AI Smooth Poly (s)

AI Assistance (a)

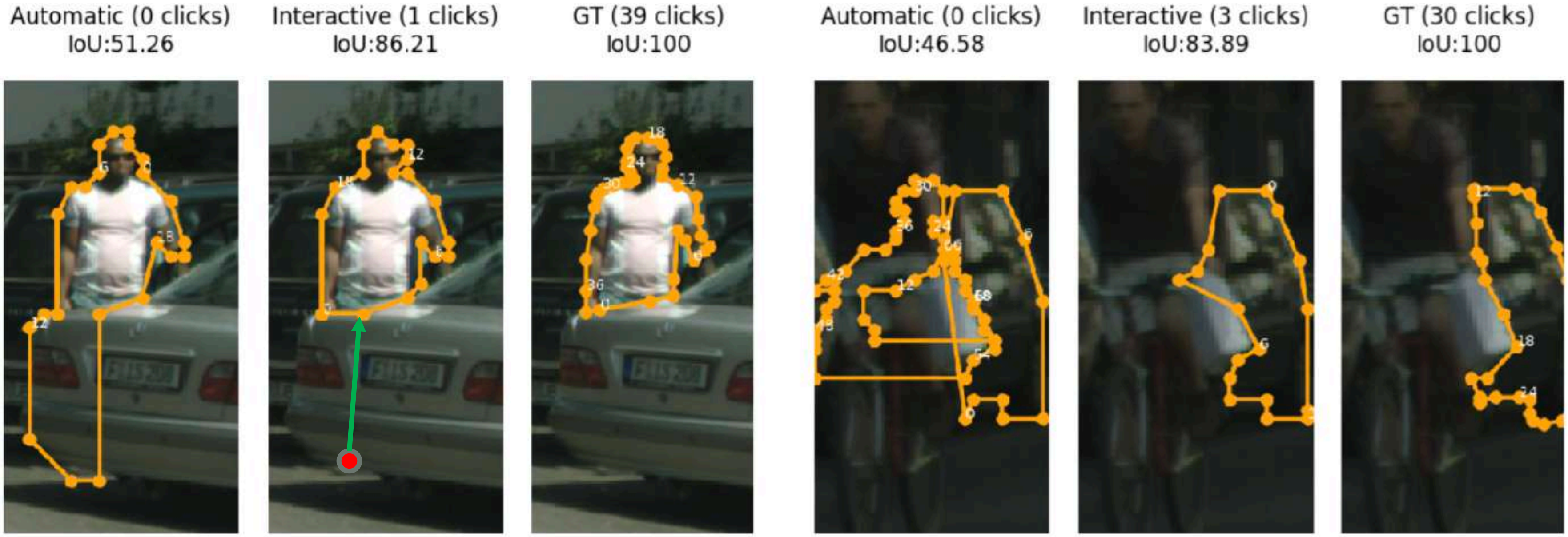
Use fine polygon (g)



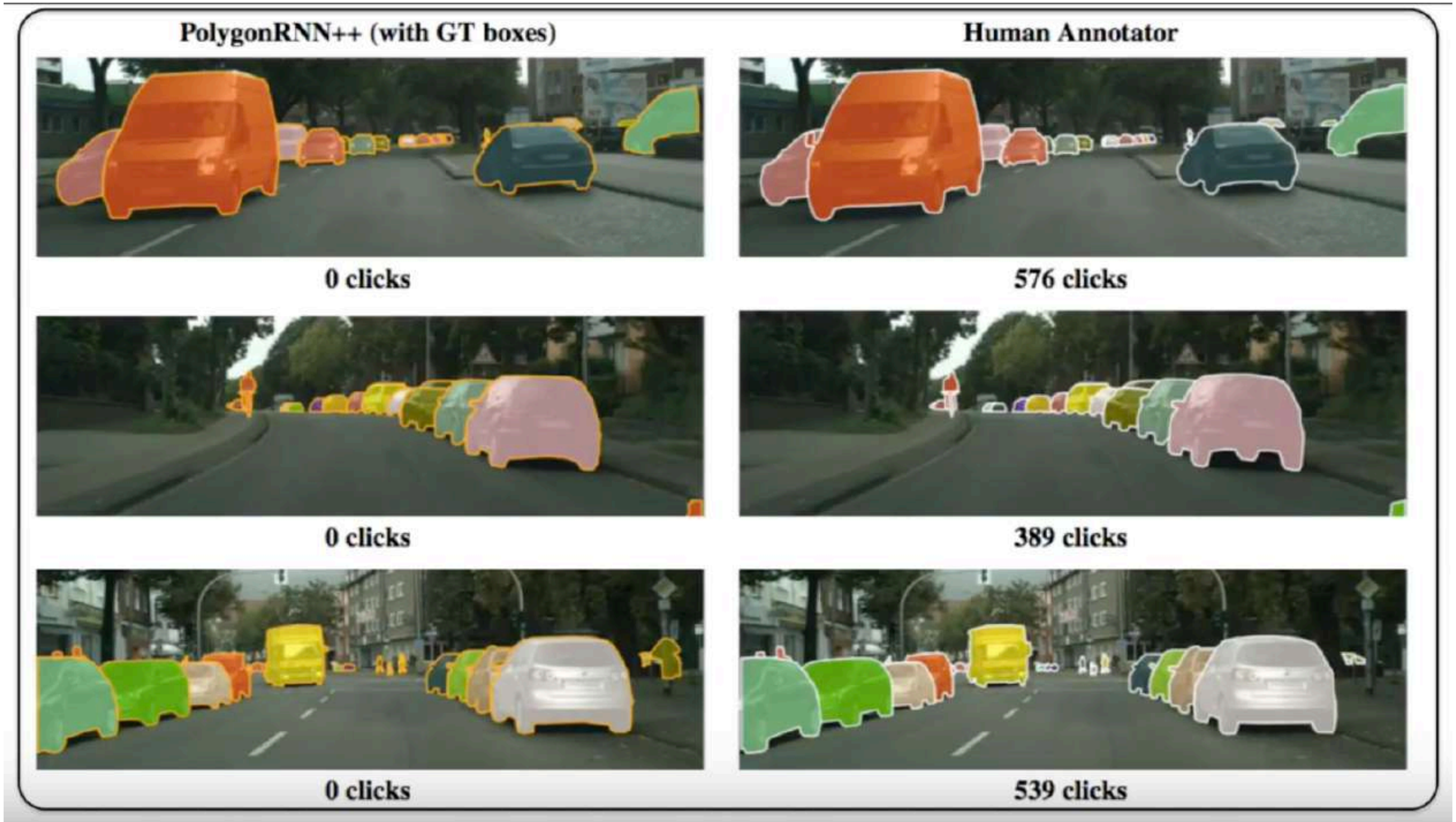
Thanks to NVIDIA for providing support in the form of GPUs for this demo.
We would also like to thank Masha Shugrina and Maxwell Fang for their help with the tool.
© University of Toronto

Polygon RNN : Bring human to the loop

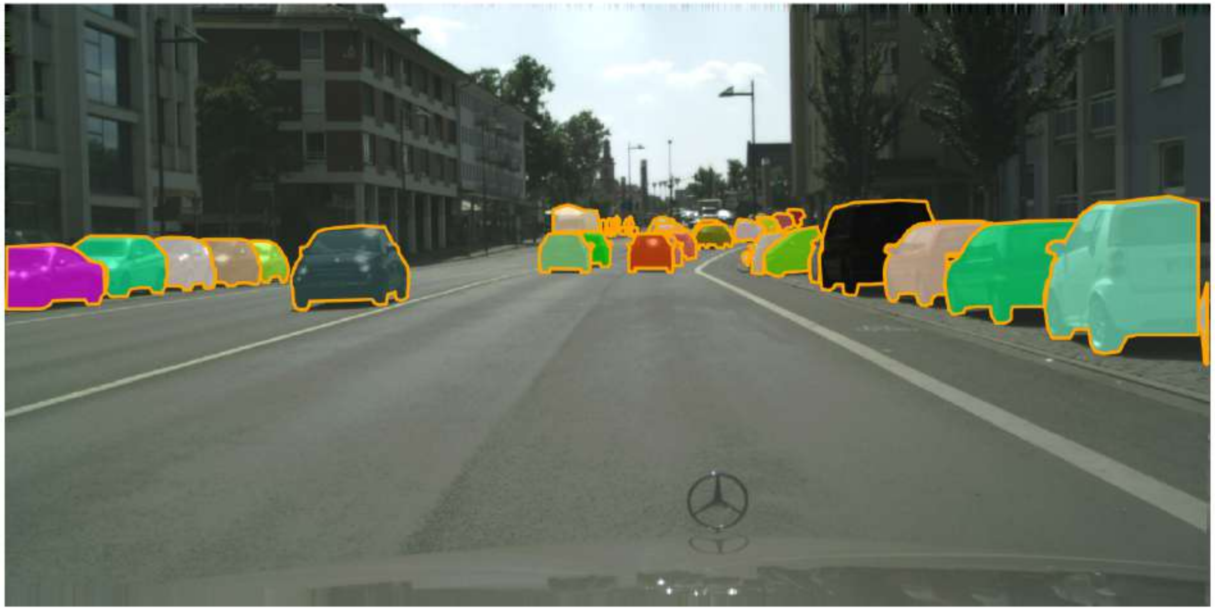
Drag key point and redo inference from that point.



Polygon RNN : Bring human to the loop



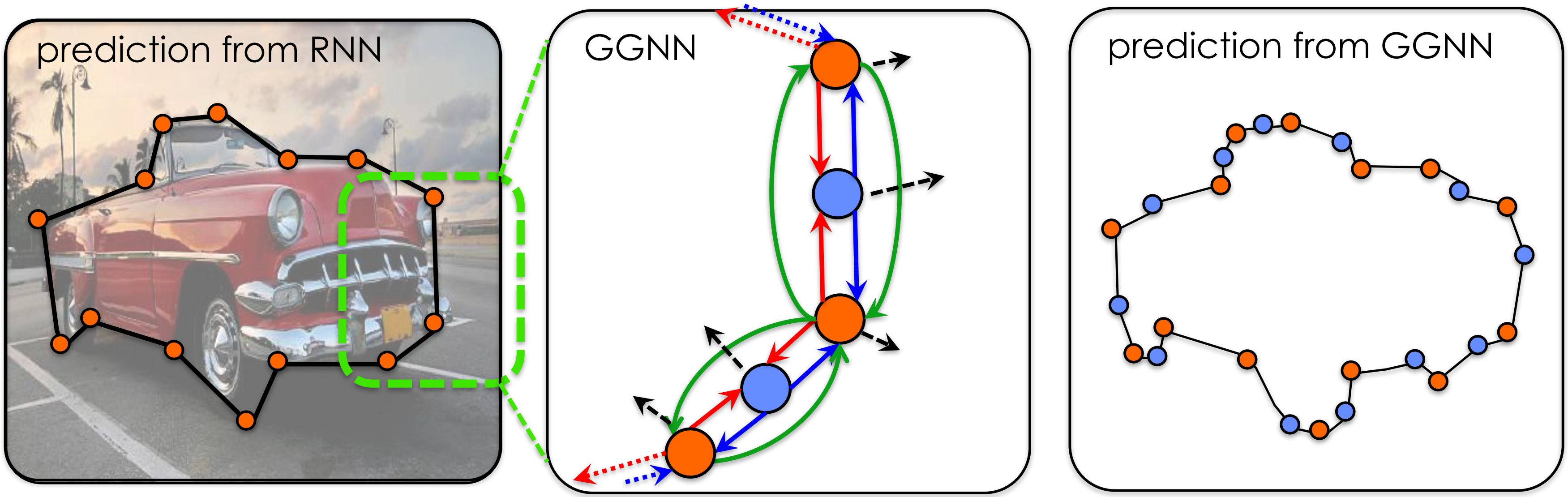
Polygon RNN++



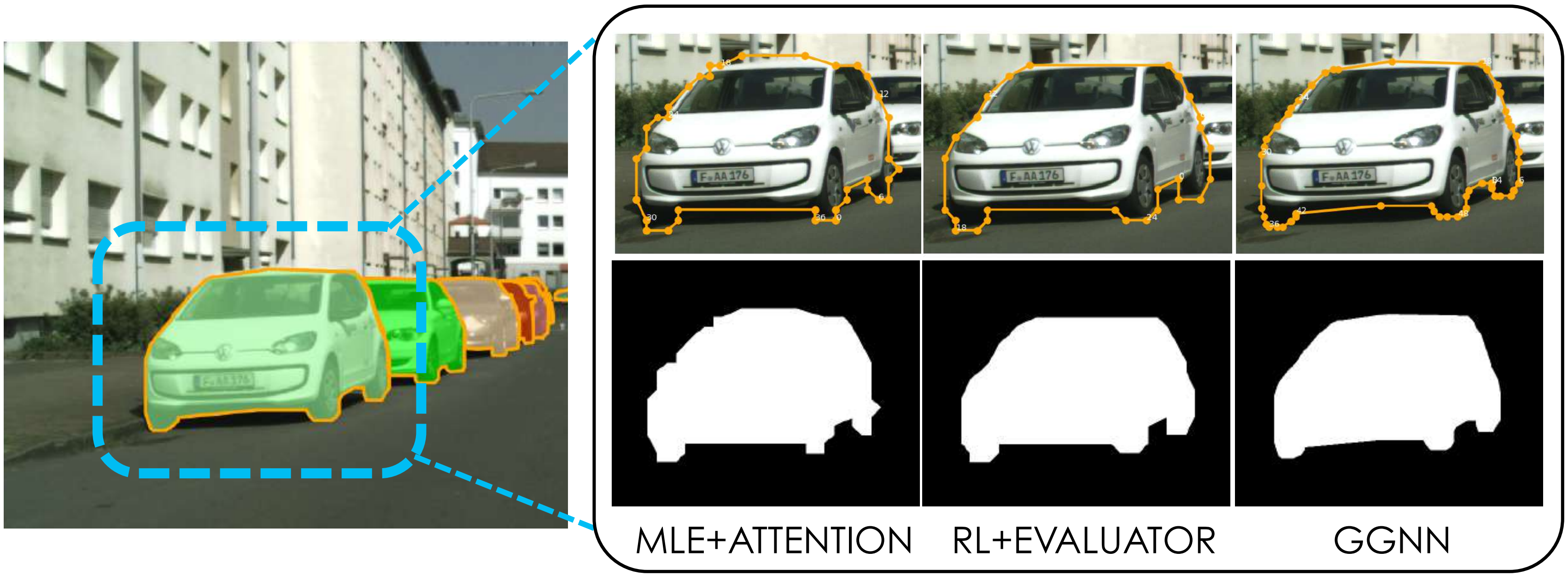
Why Polygon?

1. Human in the loop
2. **Deformable**

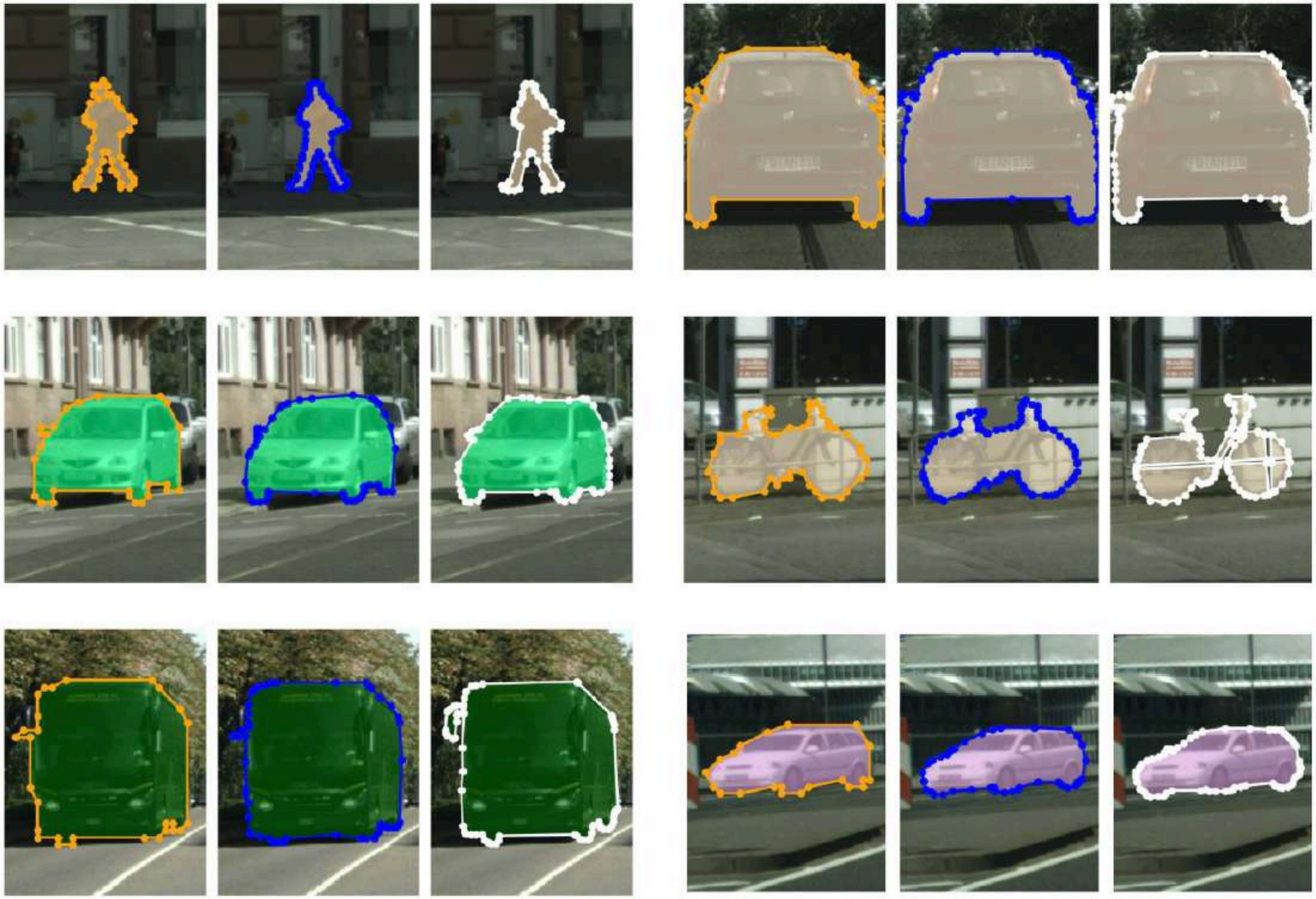
Polygon RNN++



Polygon RNN++



Polygon RNN++



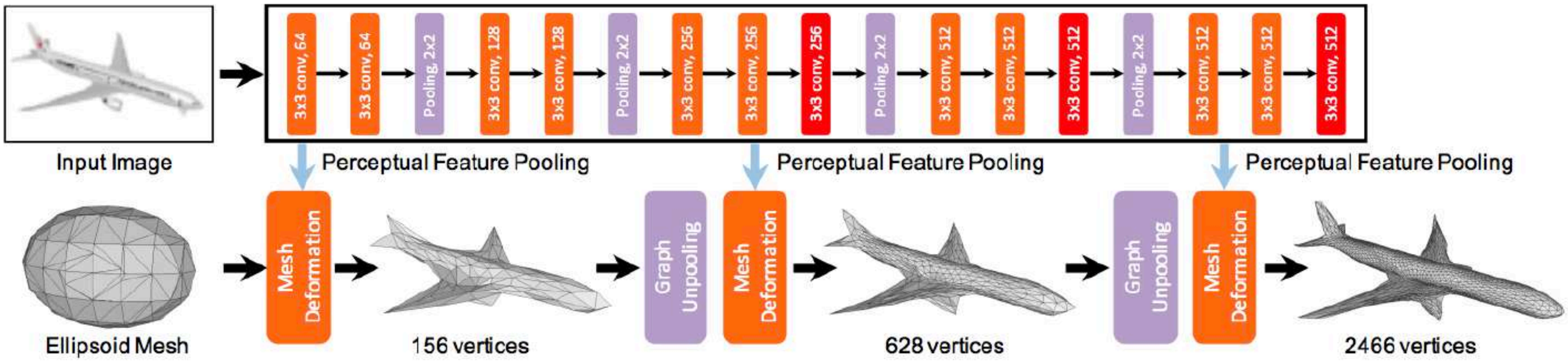
Pixel2Mesh

Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images

Nanyang Wang[1 ⋆], Yinda Zhang[2 ⋆], Zhuwen Li[3 ⋆],
Yanwei Fu[4], Wei Liu[5], Yu-Gang Jiang[1 †]

[1]Shanghai Key Lab of Intelligent Information Processing,
School of Computer Science, Fudan University
[2]Princeton University [3]Intel Labs [4]School of Data Science, Fudan University [5]Tencent AI Lab
nywang16@fudan.edu.cn yindaz@cs.princeton.edu lzhuwen@gmail.com
yanweifu@fudan.edu.cn wl2223@columbia.edu ygj@fudan.edu.cn

Pixel2Mesh



Input Image
3x3 conv, 64
3x3 conv, 64
Pooling, 2x2
3x3 conv, 128
3x3 conv, 128
Pooling, 2x2
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
Pooling, 2x2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
Pooling, 2x2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
Perceptual Feature Pooling
Perceptual Feature Pooling
Perceptual Feature Pooling
Ellipsoid Mesh
Mesh Deformation
156 vertices
Graph Unpooling
Mesh Deformation
628 vertices
Graph Unpooling
Mesh Deformation
2466 vertices

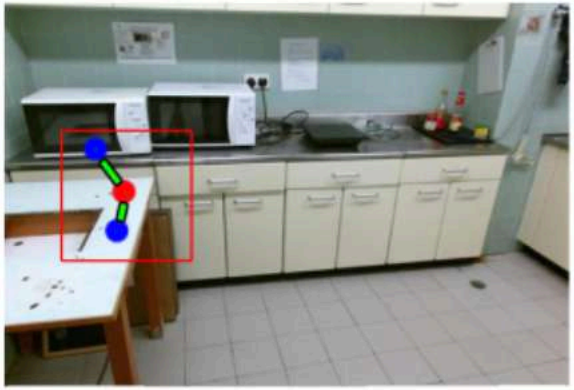
Pixel2Mesh

3D Graph Neural Networks for RGBD Semantic Segmentation

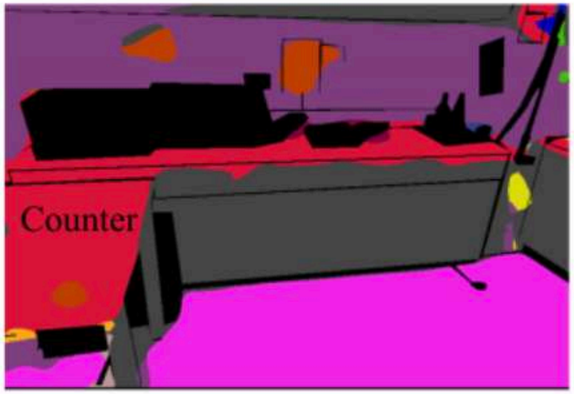
3D Graph Neural Networks for RGBD Semantic Segmentation

Xiaojuan Qi[†] Renjie Liao[‡,§] Jiaya Jia[†,♭] Sanja Fidler[‡] Raquel Urtasun[§,‡]
[†] The Chinese University of Hong Kong [‡] University of Toronto
[§] Uber Advanced Technologies Group [♭] Youtu Lab, Tencent
{xjqi, leojia}@cse.cuhk.edu.hk {rjliao, fidler, urtasun}@cs.toronto.edu

3D Graph Neural Networks for RGBD Semantic Segmentation



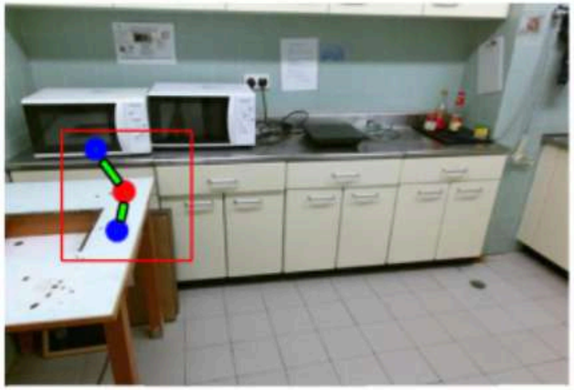
(a). 2D Image



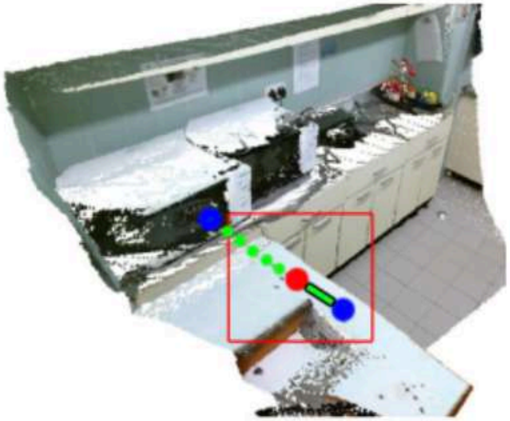
(c) 2D Prediction

1. We want to propogate information between nodes because nearby pixel are likely to be the same object.
2. We link pixels which are nearby by graph.
3. The link between top blue node and red middle node should not exist because they are actually not in the same surface.
4. How to solve this?
5. ***RGBD 3D image & graph to help***

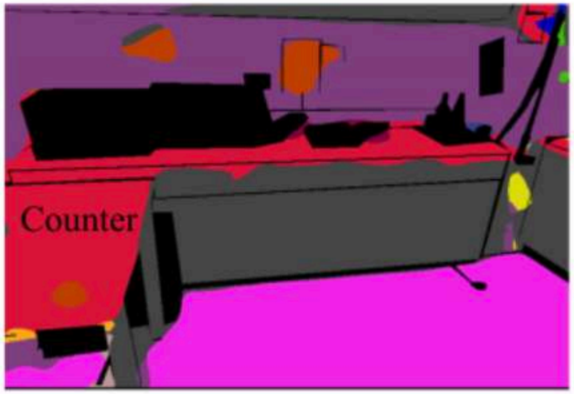
3D Graph Neural Networks for RGBD Semantic Segmentation



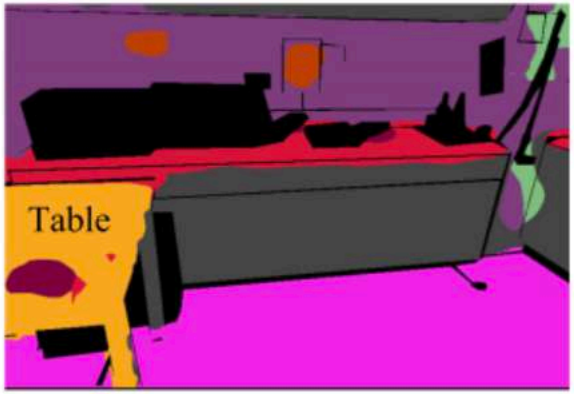
(a). 2D Image



(b) 3D Point Cloud



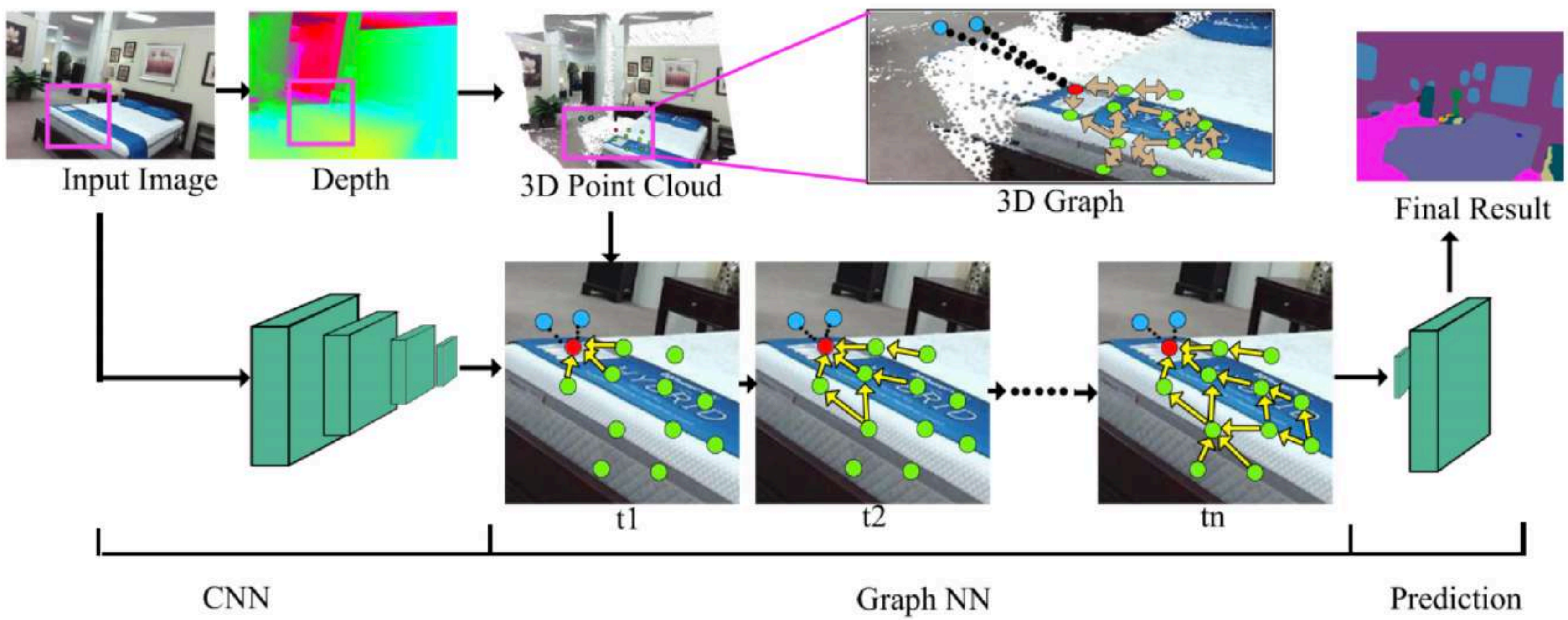
(c) 2D Prediction



(d) Ours (3D Prediction)

Dashed line means link exists in 2D image but not in 3D point cloud

3D Graph Neural Networks for RGBD Semantic Segmentation



References

1. *Michaël Defferrard, Xavier Bresson and Pierre Vandergheynst.* **Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering** https://arxiv.org/pdf/1606.09375.pdf
2. *Thomas N. Kipf and Max Welling.* **Semi-Supervised Classification with Graph Convolutional Networks** https://arxiv.org/abs/1609.02907
3. *Yujia Li, Daniel Tarlow, Marc Brockschmidt and Richard Zemel.* **Gated Graph Sequence Neural Networks** https://arxiv.org/abs/1511.05493
4. *David Acuna ,Huan Ling ,Amlan Kar and Sanja Fidler.* **Efficient Interactive Annotation of Segmentation Datasets with Polygon-RNN++** https://arxiv.org/abs/1803.09693
5. *Lluis Castrejon ,Kaustav Kundu ,Raquel Urtasun and Sanja Fidler.* **Annotating Object Instances with a Polygon-RNN** https://www.cs.toronto.edu/~urtasun/publications/castrejon_etal_cvpr17.pdf
6. *Nanyang Wang , Yinda Zhang , Zhuwen Li , Yanwei Fu , Wei Liu and Yu-Gang Jiang.* ***Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images*** https://arxiv.org/pdf/1804.01654.pdf
7. *Xiaojuan Qi, Renjie Liao, Jiaya Jia, Sanja Fidler and Raquel Urtasun.* ***3D Graph Neural Networks for RGBD Semantic Segmentation*** http://www.cs.toronto.edu/~rjliao/papers/iccv_2017_3DGNN.pdf