

Reinforcement Learning with Unity 3D: Autonomous Garbage Collector

Sangram Gupta Damian Bogunowicz HyunJun Jung





OKTOBERFEST!!!



1 billion euros

value of the Oktoberfest to the Munich economy

6 million

visitors celebrate Oktoberfest in Munich every year

1000 tons

total amount of waste produced at Oktoberfest



- **Massive events**
- **Large scale operation**
- **Functional 24/7**
- **Autonomous, intelligent**



Damian

Passionate about making machines autonomous and intelligent.



HyunJun

Biomedical Computing student, loves Computer Vision and Deep learning.



Sangram

Exploring new technologies in Computer Vision and also into getting decent grades.

G.E.A.R

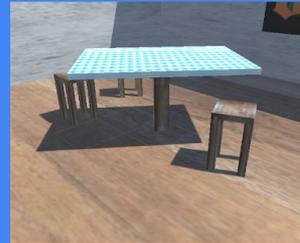
Garbage Evaporating Autonomous Robot



Environment

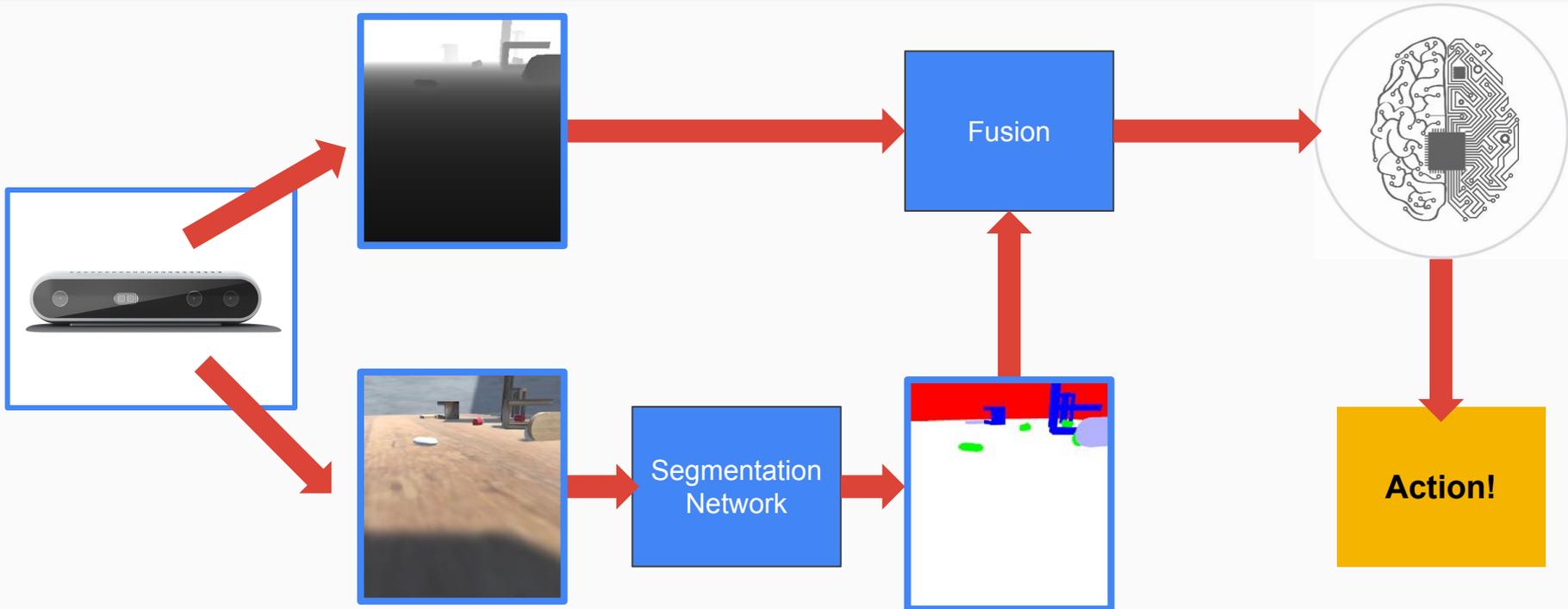


Collect!

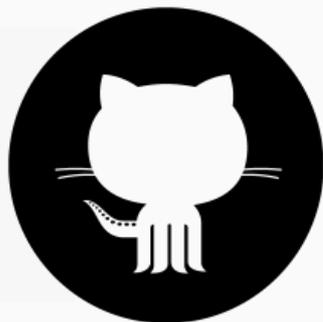


Avoid!

Perception, Cognition, Action



Software

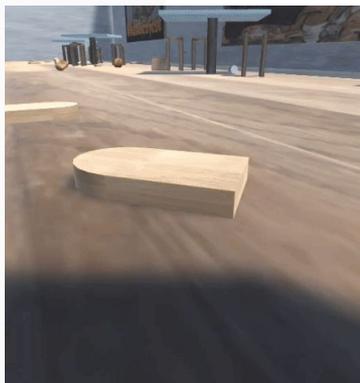
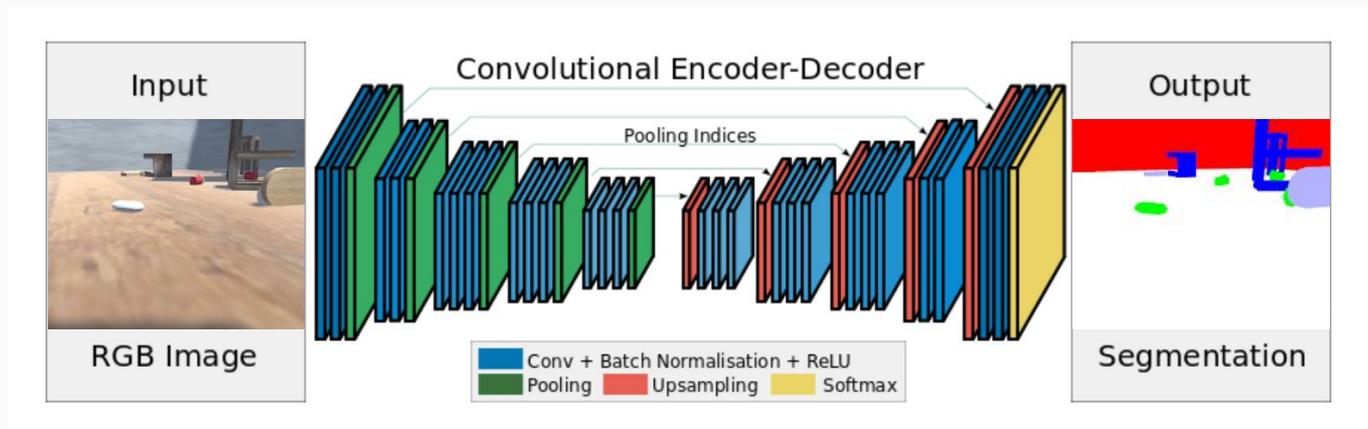


<https://github.com/GeorgeSeif/Semantic-Segmentation-Suite>

Algorithm

- **Semantic Segmentation**
(SegNet, Badrinarayan et al., 2015
<https://arxiv.org/pdf/1511.00561.pdf>)
- **Behavioral Cloning**
(Bain and Sommut, 1999
<https://www.ijcai.org/proceedings/2018/0687.pdf>)
- **Proximal Policy Optimization**
(Schulman et al., 2017
<https://arxiv.org/abs/1707.06347>)
- **Our own heuristic**

Semantic Segmentation



input



prediction



ground truth

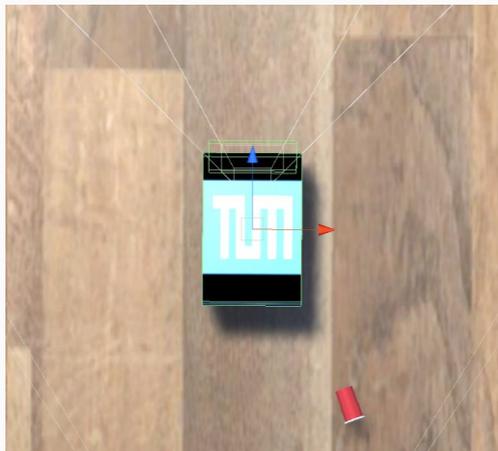


Actions

➡️ ⬅️ ✖️ Left/Right/Empty

⬆️ ⬆️ ✖️ Forward/Backward/Empty

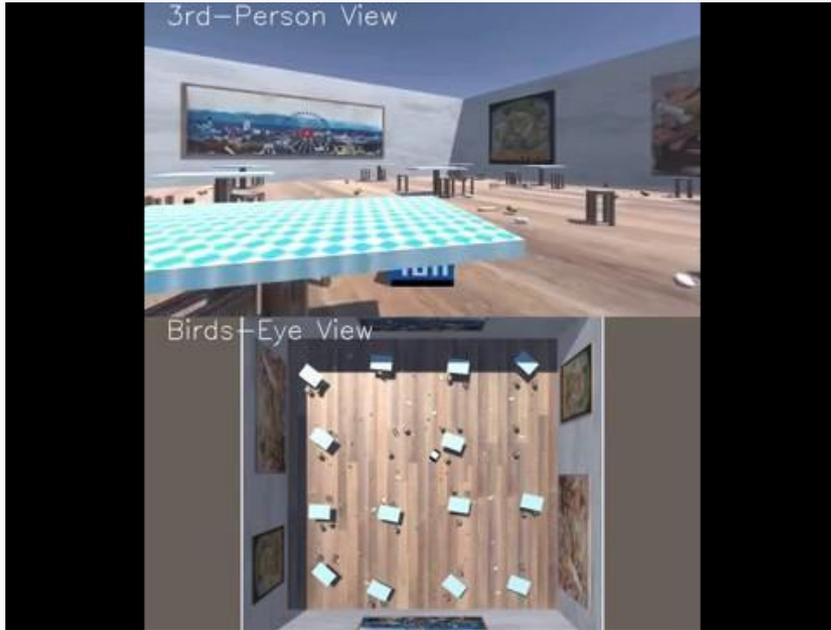
🟢 🚫 Grabber On/Grabber Off



Rewards

- Collect non-collectible item
- Slam against the wall
- Slam against the obstacle
- Punishment per step
- Punishment per grabber activation
- Reward for forward movement
- Collect garbage

Behavioral Cloning

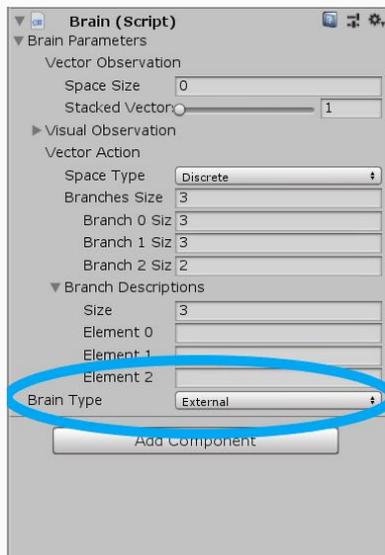


- Short training time
- Only as clever as human player
- Good for naive agents

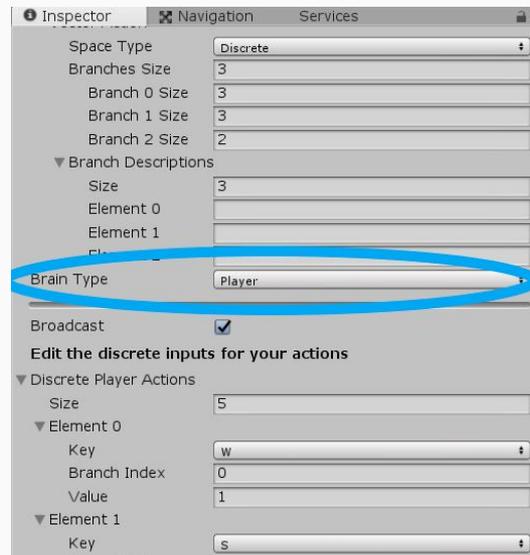
Behavioral Cloning



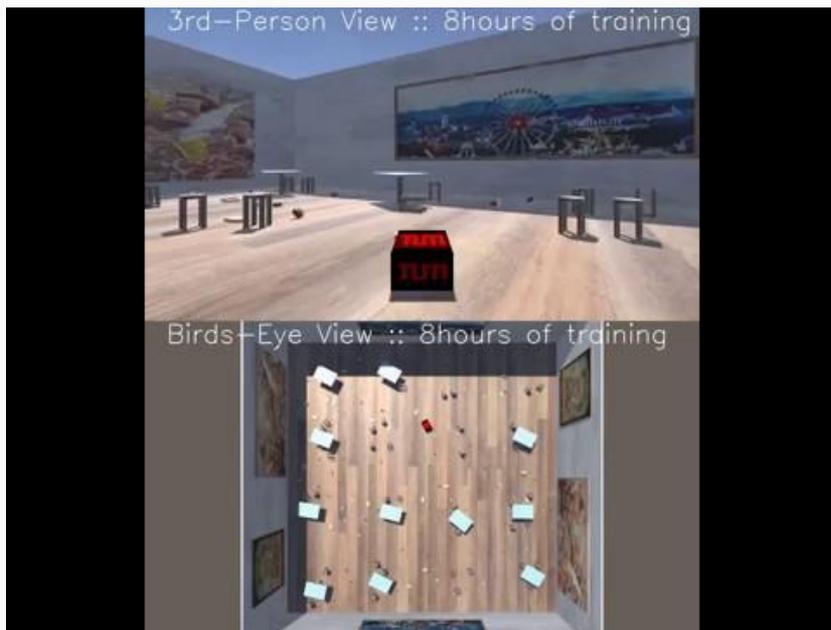
Student Brain



Teacher Brain

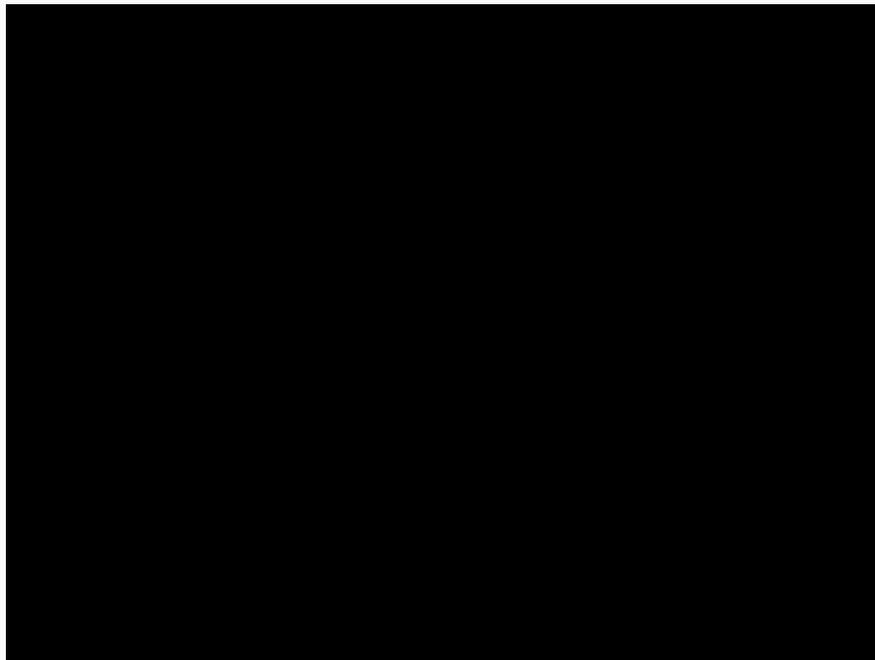


PPO: Single-Agent



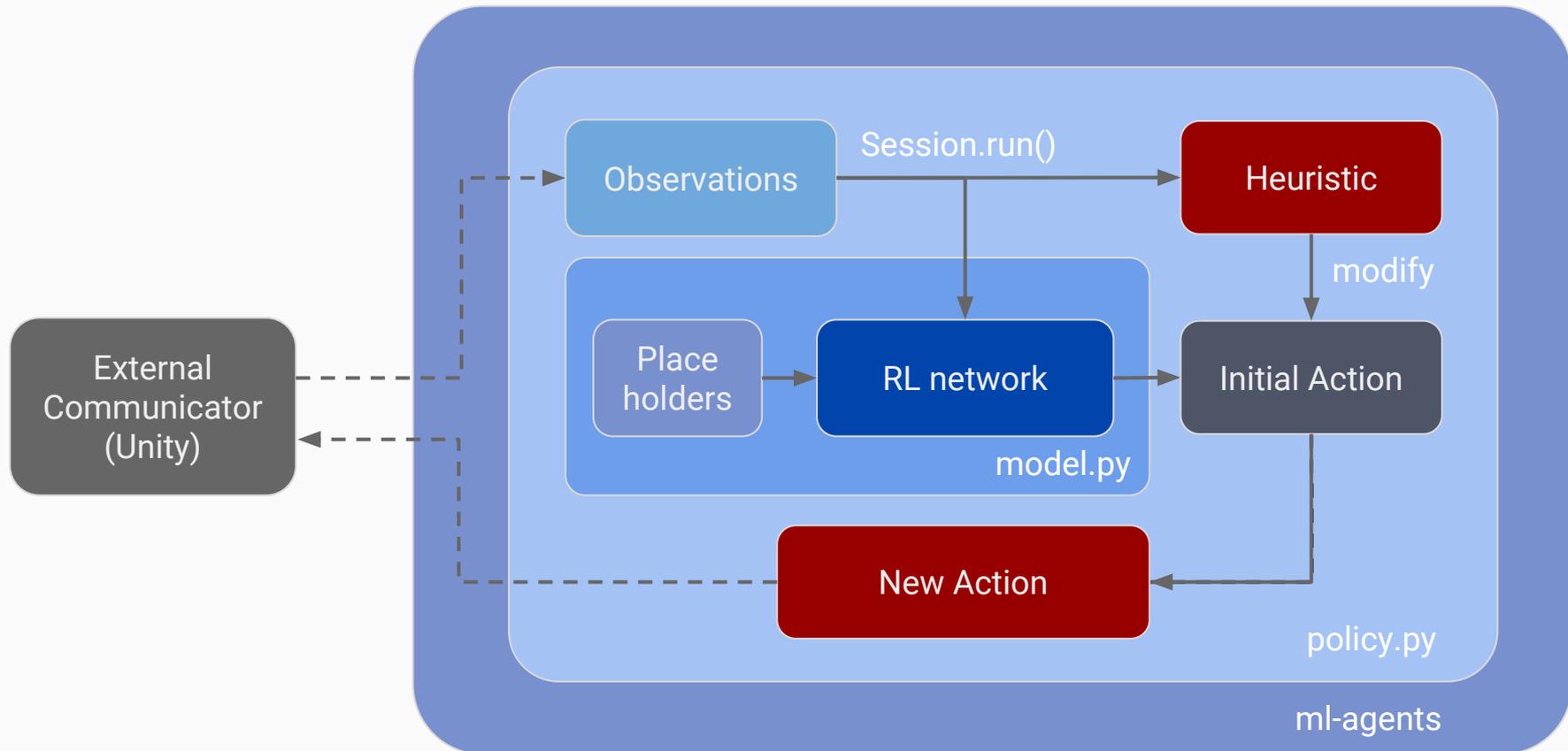
- Long training time
- Increase punishments slowly
- About 40h of training
- Great learning experience!

PPO: Heuristic

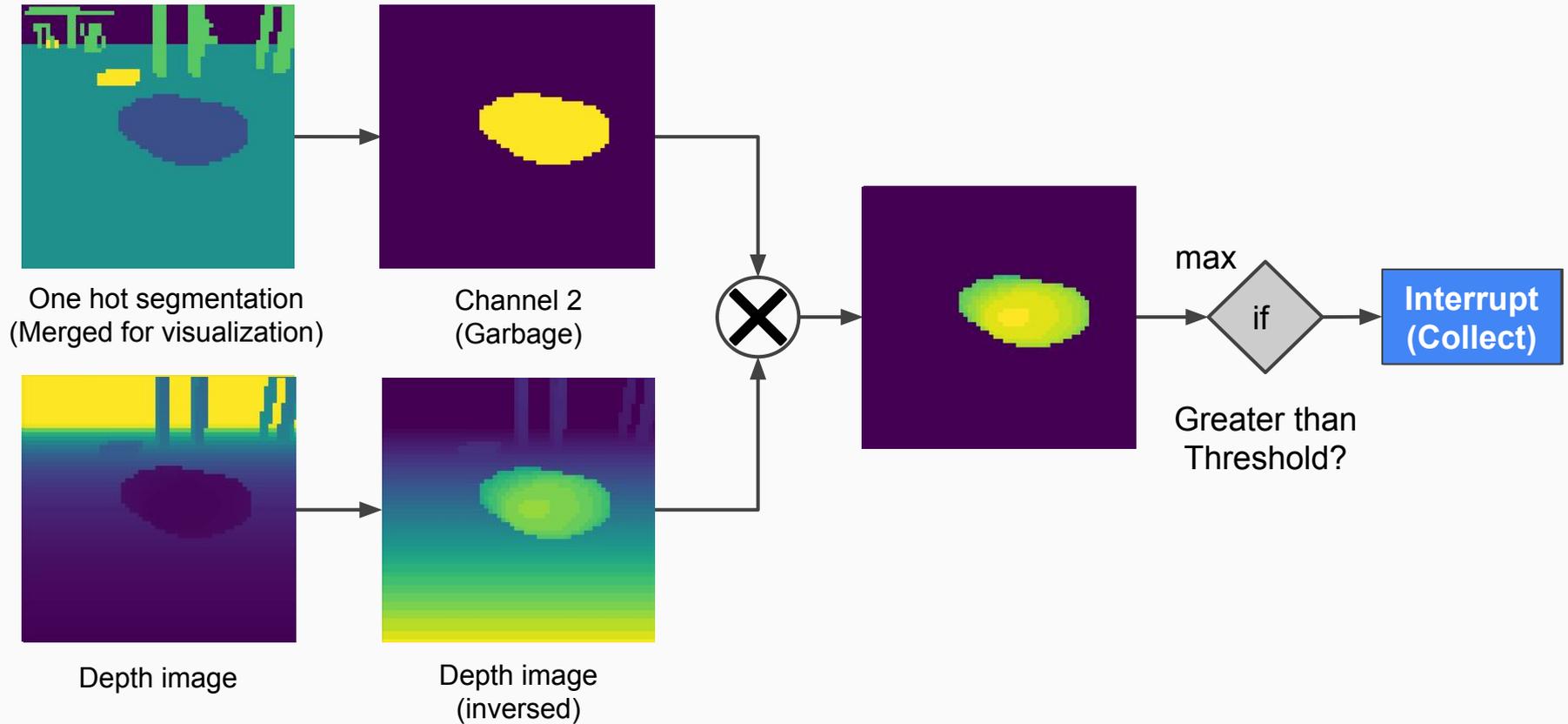


- PPO for navigation and heuristic for collection
- Feasible for simple action
- Medium training time

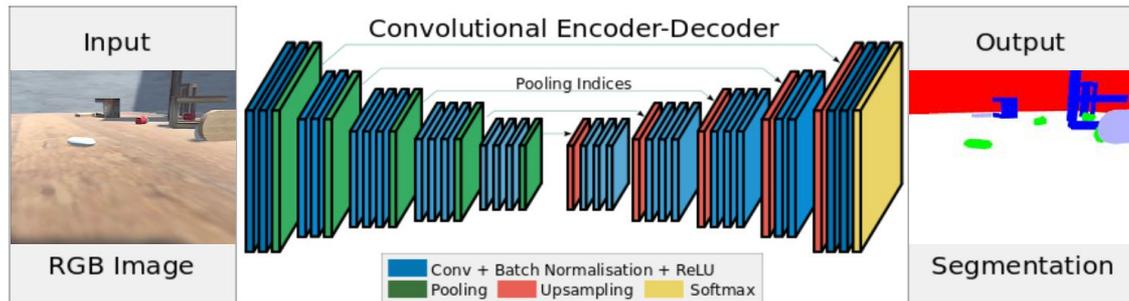
Heuristics : API Perspective



Heuristics : Algorithm



PPO: with SegNet



Two Approaches:

1. Train PPO with SegNet

- Easiest way to implement
- It takes about 5s to generate an observation

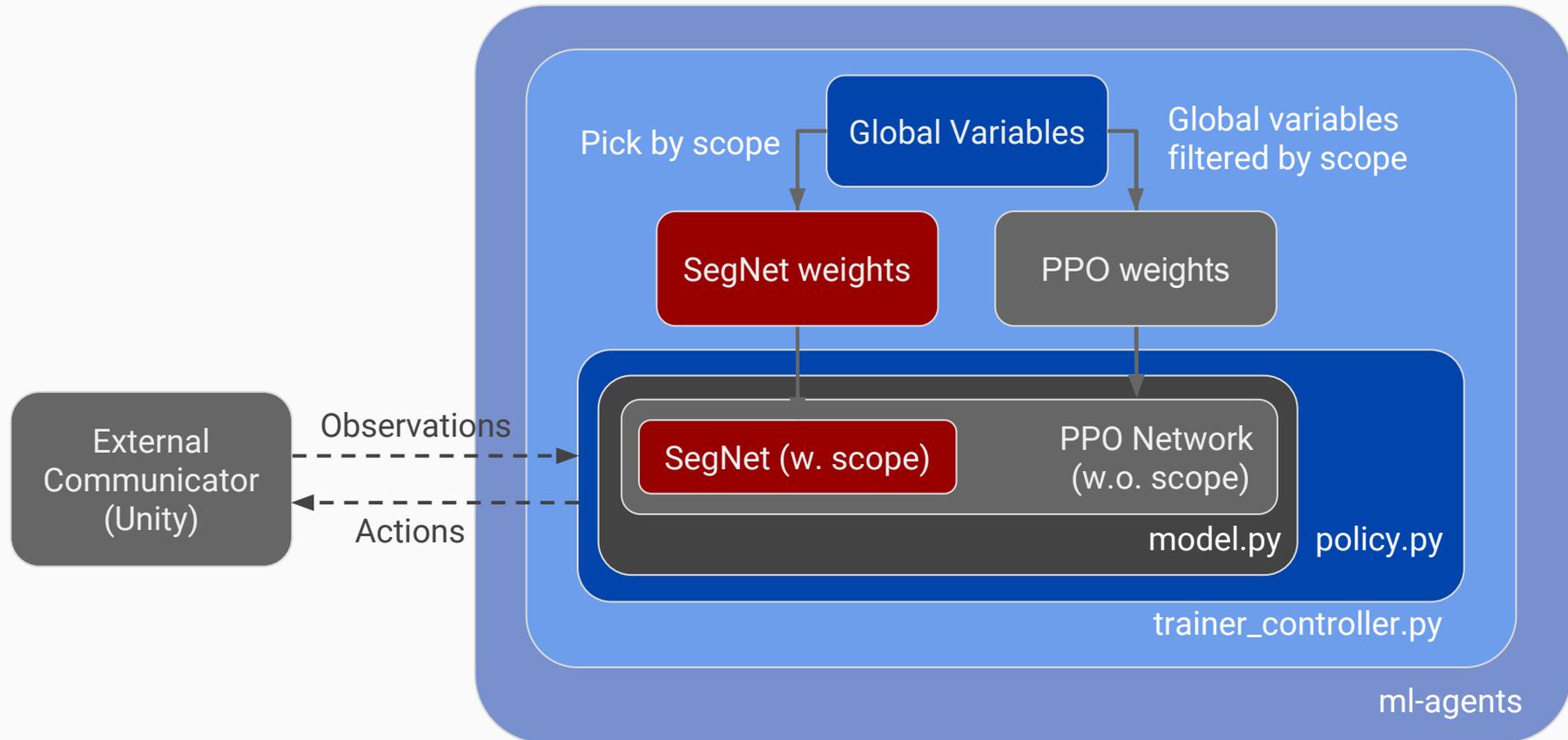


2. Train PPO network separately

- Combine two only in test time
- Tricky to implement
- No effect on performance during training



PPO + SegNet : API perspective



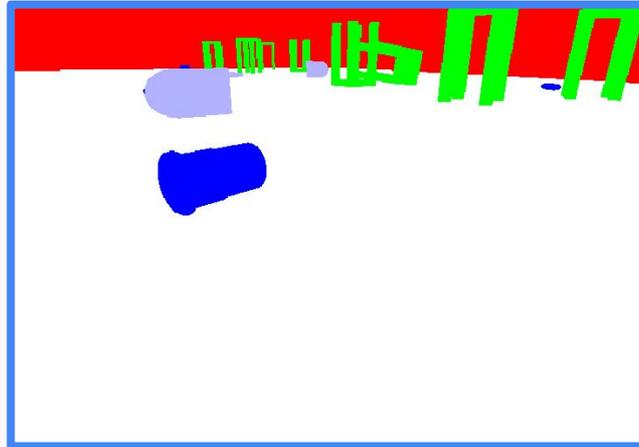
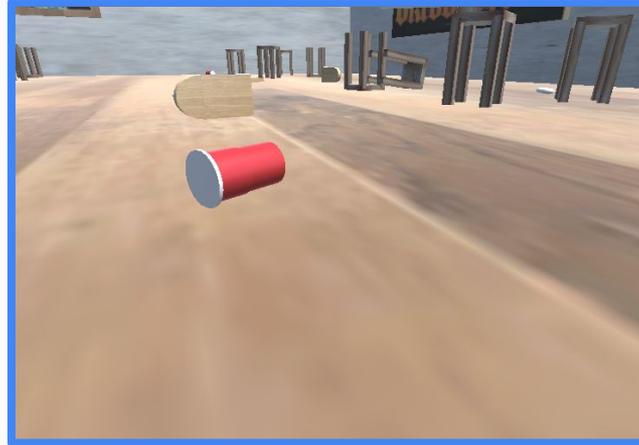
SegNet In Action:



- Computationally expensive
- Reflects real world implementation (RealSense camera)
- Easy modification of its objective

A simple modification :

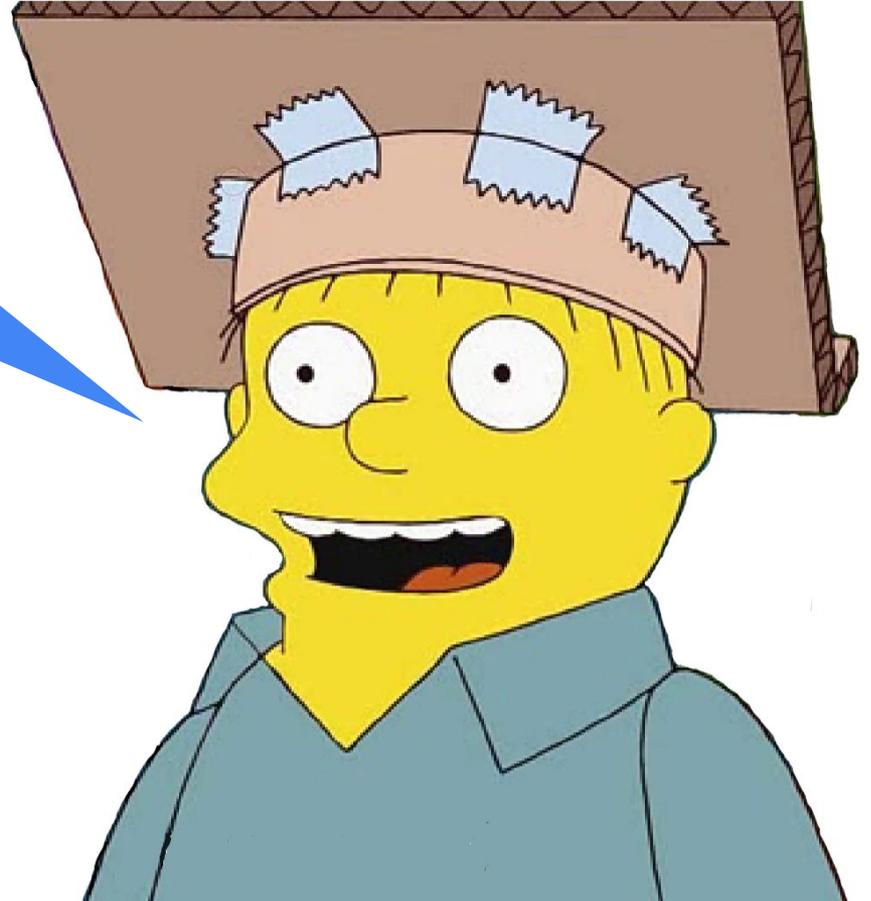
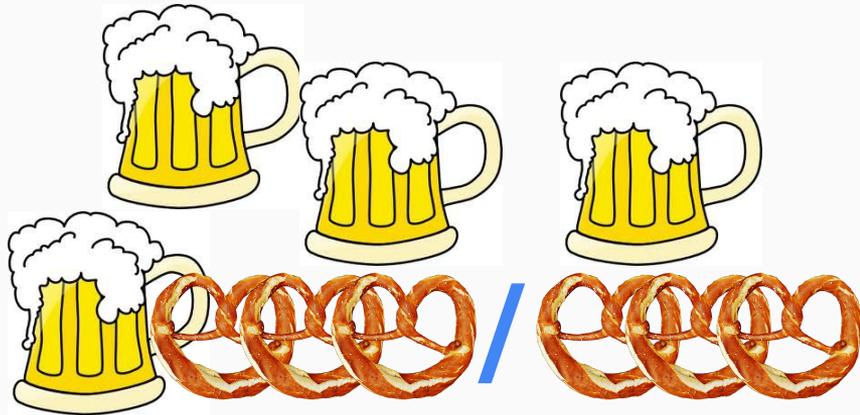
I need to collect channel 2
I should not collect channel 5
I have to avoid channel 1,4



-  Channel 1
-  Channel 2
-  Channel 3
-  Channel 4
-  Channel 5

Plot twist : The Furniture Collector

I collect
furniture now!



The Furniture Collector In Action :



Room for improvement

- **Install the actual mechanism for garbage collection**
- **Deploy the algorithm on machine can handle real-time semantic segmentation**
- **Transfer the knowledge from simulation to a real robot with RealSense camera**
- **Make the world a better and cleaner place!**

Outlook for the future: fleet of autonomous robots



Thank You For Attention!



Image references (in order)

- <https://www.euronews.com/2018/09/22/it-s-tapped-octoberfest-kicks-off-in-munich>
- <https://www.abendzeitung-muenchen.de/inhalt.wiesn-nachbarn-in-sorge-oktoberfest-muell-urin-und-erbrochenes-ob-diese-hotline-helfen-kann.6a6cb3f8-06f5-419b-bbdf-4d8324707bd0.html>
- <https://www.dw.com/en/earth-lovers-in-lederhosen-oktoberfest-goes-green/a-18722603>
- <https://imgur.com/gallery/lqYpC>
- <https://www.desicomments.com/desi/cartoons/homer-simpson/>