International Advisory Board Meeting for
Computer Engineering Industry 4.0 DOE Project
U. Verona, Italy, May 25, 2023

"Edge-Brain" :

Edge Computing AI Solutions for Easy-to-Upgrade of Flexibility in Smart Manufacturing

Prof. Sukhan Lee
Sungkyunkwan University
School of Artificial Intelligence

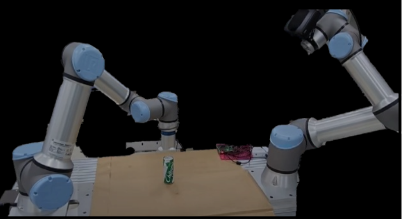
1

Table of Contents

- Introduction
- Scenarios
- AI-Vision Solutions
  - Technologies
  - Applications
- AI-Robot Skills
  - Unit Task Skills
  - Industrial Bin Picking
  - Industrial Bin Packing

2

Introduction



3

Introduction

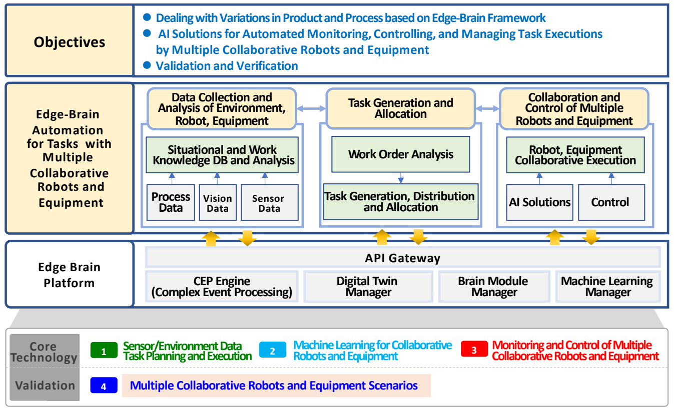
Background

- Flexibility: Dealing with Variations in Products and Processes with Minimum Effort and Time Based on Edge-Brain Framework .
- Incorporation of Edge Computing AI Solutions for AI-Vision Capabilities, Flexible Gripping, Multiple Collaborative Robots and Equipment into Automated Planning and Scheduling as well as Monitoring, Control and Execution.
- Building an Easy-to-Implement Environment over Ultra Low Latency Networks for integrating Edge Computing AI Solutions with Legacy Systems.
- Project by an Industry-Government Research Institute-University Consortium

4

Introduction

Edge Brain Framework

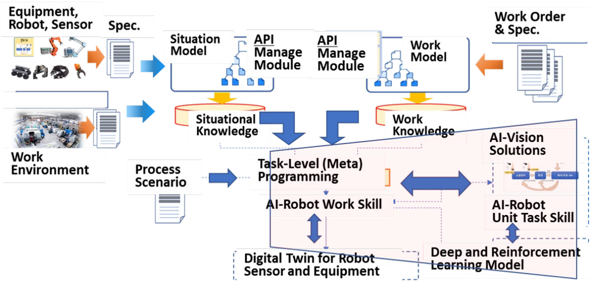


5

Introduction

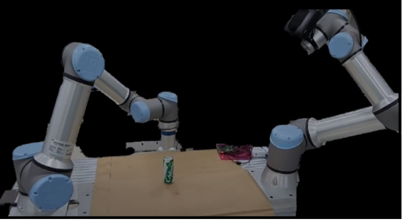
Edge Computing AI Solution

- **Panoptic Segmentation**: Semantic and Instance Segmentation
- **Cascaded Yolo-YOLACT with Depth-Based Tone Mapping and Part Configuration Loss for High Performance**



6

Scenarios



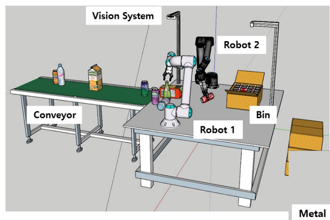
7

Scenarios

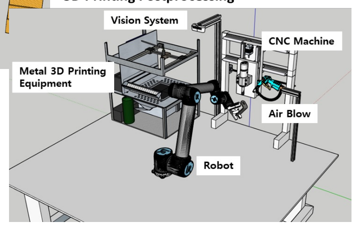
Application Scenarios: Example 1

- **Multi-Robot Cooperative Drinks Bin Packing**
- **Robot-CNC Machine Collaborative 3D Printing Postprocessing**

Multi-Robot Cooperative Drinks Bin Packing



Robot-CNC Machine Collaborative 3D Printing Postprocessing



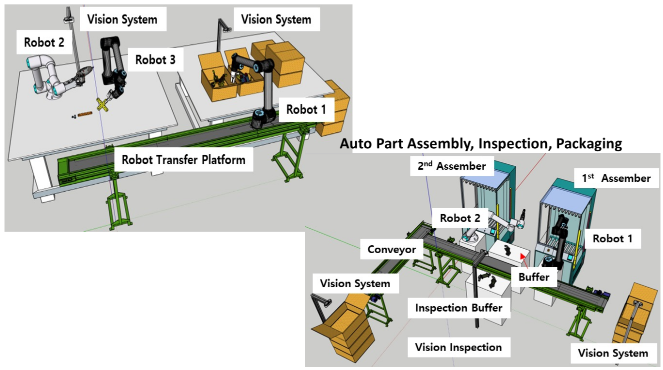
8

Application Scenarios: Example 2

Scenarios

- Multi-Robot Collaborative Toy-Kit Assembly
- Multi-Robot Auto Part Assembly, Inspection and Packaging

Multi-Robot Toy Assembly

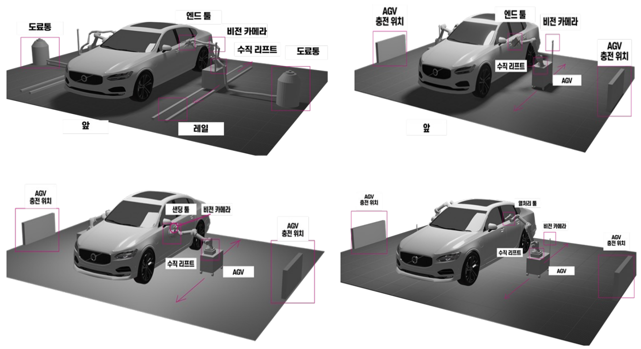


9

Application Scenarios: Example 3

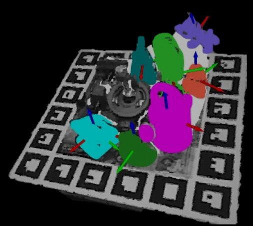
Scenarios

- Robotic Auto-Body Sanding, Painting, Surface Treatment



10

AI-Vision Solutions: Technologies

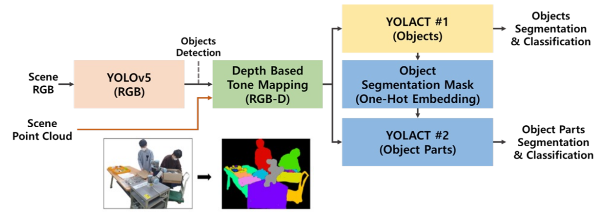


11

3D Workspace Modeling: Real-Time Panoptic Segmentation

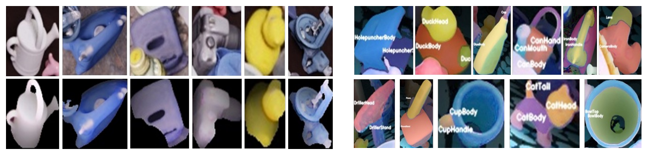
AI-Vision Solutions: Technologies

- Panoptic Segmentation: Semantic and Instance Segmentation
- Cascaded Yolo-YOLACT with Depth-Based Tone Mapping and Part Configuration Loss for High Performance



Object Segmentation

Part Segmentation

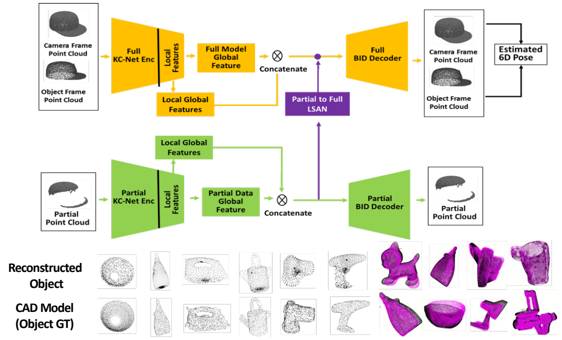


12

AI-Vision Solutions: Technologies

3D Workspace Modeling: Real-Time 6D Pose Estimation

- Fast and Accurate 6D Pose Estimation between Camera and Object Frames
- Dual Associative Point Auto-Encoder for 6D Pose Estimation with Full Model Reconstruction

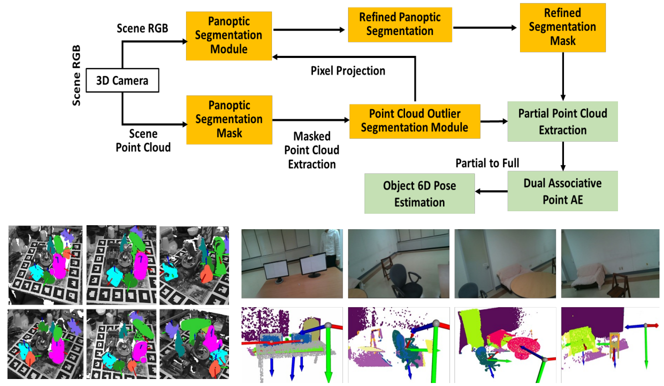


13

AI-Vision Solutions: Technologies

3D Workspace Modeling

- Real-Time 3D Workspace Modeling System with Integrated Panoptic Segmentation and 6D Pose Estimation



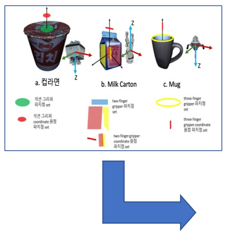
14

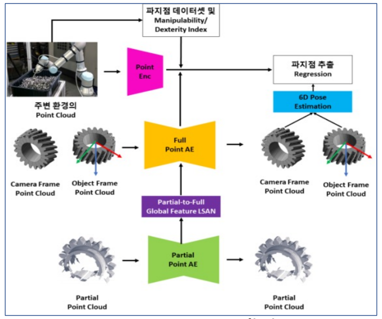
AI-Vision Solutions: Technologies

Optimal Grasp Pose Prediction Based on AI-Vision Solutions

- Collision-Avoidance Optimal Grasp Pose Prediction Based on Object Model Reconstruction, 6D Pose Estimation and 3D Workspace Modeling

6D Graspable Pose Manifold



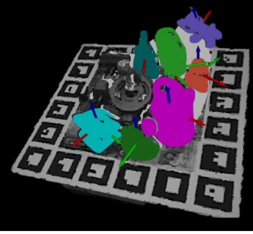


Deep Grasp Pose Prediction

15

AI-Vision Solutions:

Applications

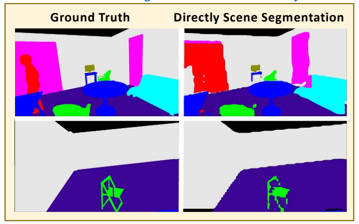


16

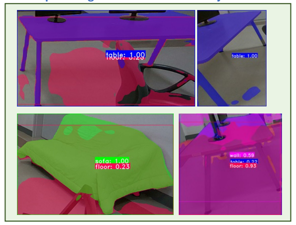
Panoptic Segmentation

- Benefits of a Cascaded YOLO-YOLACT (mIoU 90.47%) structure
  - ✓ High Speed: over 20 FPS/Scene
  - ✓ Accurate Boundary Segmentation
  - ✓ Reliable Segmentation of Small Objects
  - ✓ High Accuracy in Part Segmentation by Preventing Inter-Class Confusion
  - ✓ mIoU 90.47% with Cascaded YOLO-YOLACT compared to mIoU 72.51%

Ablation Study 1:
Perform scene level segmentation instead of object level

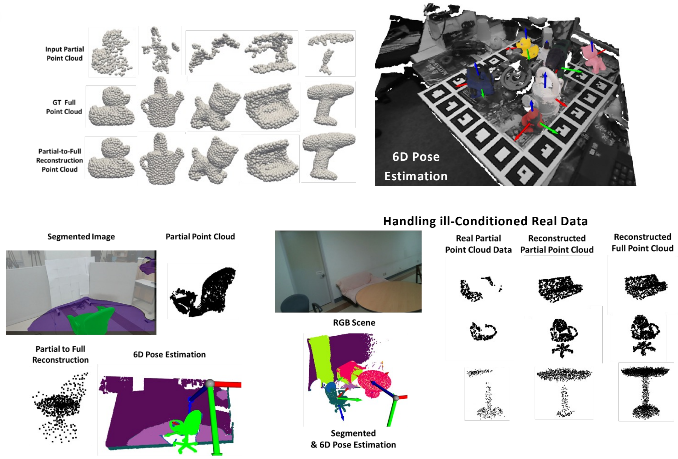


Ablation Study 2:
Not separating models based on object features



17

6D Pose Estimation

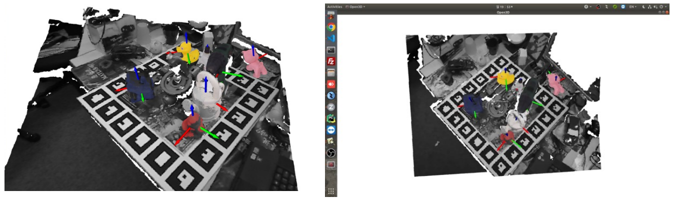


AI-Vision Solutions: Applications

18

LineMod Competition

- LineMod Dataset: Dataset for comparing 6D pose estimation accuracy in real environments with heavily occluded objects.



- Performance Comparison in the LineMod Leader Board

Rank		AR	AR(VSD)	AR(MSSD)	AR(MSPD)	Processing Time
1	PFA-MixPBR	0.797	0.658	0.843	0.890	1.7 sec.
2	Ours	0.792	0.719	0.825	0.840	0.6 sec.
3	GDRNPP	0.792	0.651	0.836	0.889	
4	SurfEmb-PBR-R GBD	0.760	0.615	0.809	0.856	
5	RCVPose	0.749	0.682	0.773	0.792	

AI-Vision Solutions: Applications

19

Pick and Place Application

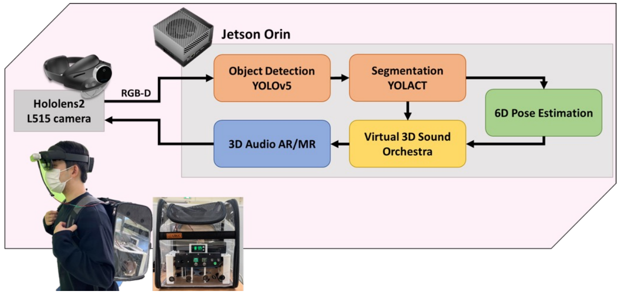
AI-Vision Solutions: Applications

20

AI-Vision Solutions: Applications

Multi-Modal 3D Workspace Modeling 1

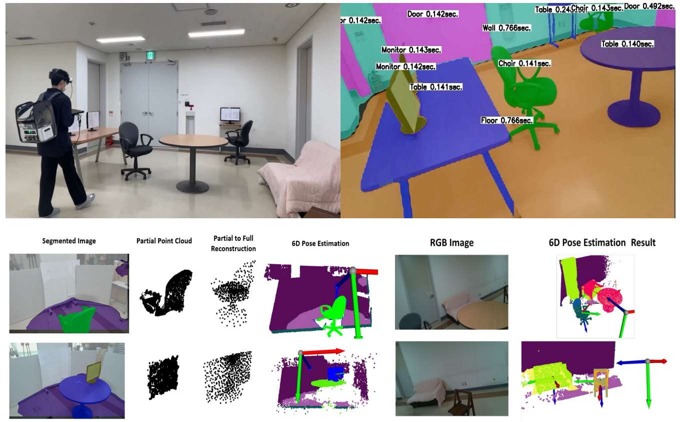
- **Deep Learning Based Visual-Audio 3D Modeling: a deep learning system implemented in a wearable device enables multi-modal 3D modeling of the surrounding environment, which is useful for the visually impaired.**



21

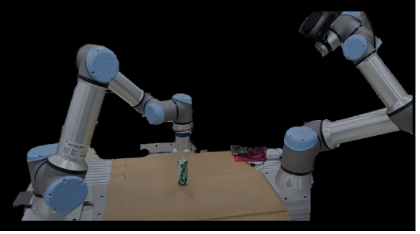
AI-Vision Solutions: Applications

Multi-Modal 3D Workspace Modeling 1



22

AI-Robot Skills



23

AI-Robot Skills

AI-Robot Skills for Automated Task-Level Programming of Collaborative Robots and Equipment

- **AI-Robot Skills combined with Conventional Engineering Solutions for ease in Implementation → A Framework of Reinforcement Learning with Sim2Real Transfer Learning**
- **Synthesis of Work Skills Based on Unit Task Skills and AI-Vision Solutions**

AI-Vision Solutions
- Object Recognition
- 6D Pose Estimation
- 3D Workspace Modeling
- Real-time Monitoring

➡

Work Skills
- Pick & Place
- Dual Robot Assembly
- Collaborative Bin Picking
- Collaborative Bin Packing

⬅

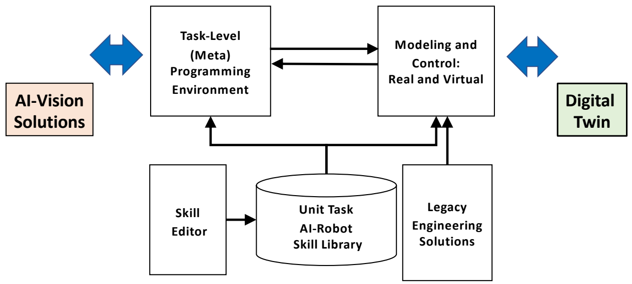
Unit Task Skills
- Collision-Free Path and Trajectory
- Pose Alignment
- Collaborative Moving and Pushing
- Optimal Grasping
- Insertion/ Screwing/Cutting

24

AI-Robot Skills

Skill-Based Task-Level (Meta) Programming Environment

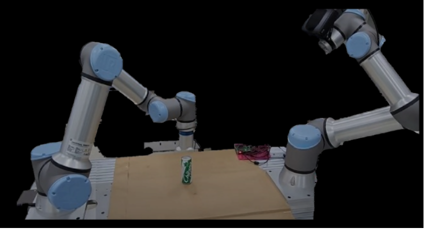
- Provision of a Unit Task Skill Library
- Skill-Based Task-Level (Meta) Programming Environment for Multiple Collaborative Robots and Equipment
- S/W Development Platform for integrating AI-Robot Skills with Conventional Engineering Solutions



25

AI-Robot Skills:

Unit Task Skills

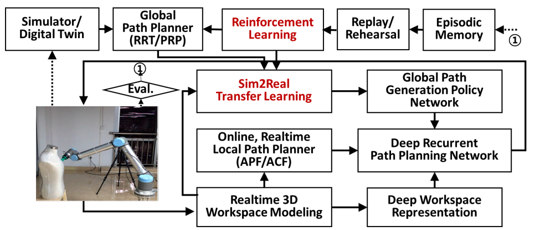


26

AI-Robot Skills: Unit Task Skills

Optimal Collision-Free Path and Trajectory Generation Skill

- Simulation/Digital Twin-Based Reinforcement Learning (RL) for Global Path/Trajectory Planning
- Sim2Real Transfer Learning with Episodic Memory to fuse RL-based Global Planning and Reactive Online Real-Time Local Planning



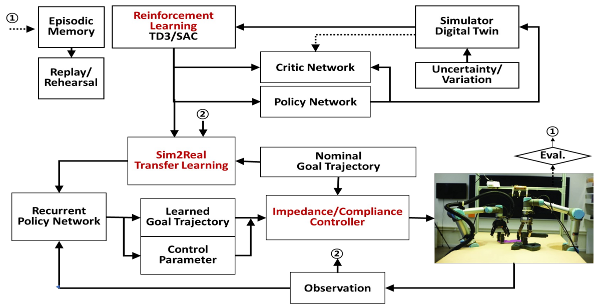
RRT: Randomly-exploring Random Tree
PRP: Personalized Route Planning
APF: Artificial Potential Field
ACF: Artificial Circular Field

27

AI-Robot Skills: Unit Task Skills

Insertion/Screwing/Cutting Skills

- Integration of Simulation-Based Reinforcement Learning and Classical Impedance/Compliance Control Based on Sim2Real Transfer Learning



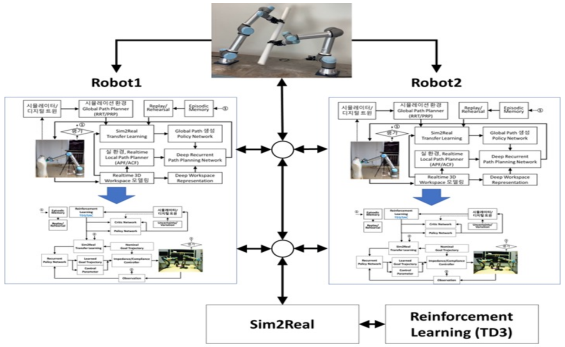
TD3: Twin-Delayed Deep Deterministic Policy Gradient, SAC: Soft Actor-Critic

28

AI-Robot Skills: Unit Task Skills

Synthesis of Multi-Robot Collaborative Work Skills Based on Unit Task Skills

- Learning Interaction Policy at the Upper Hierarchy of Skills Based on RL and Sim2Real Fusion Framework



29

AI-Robot Skills:

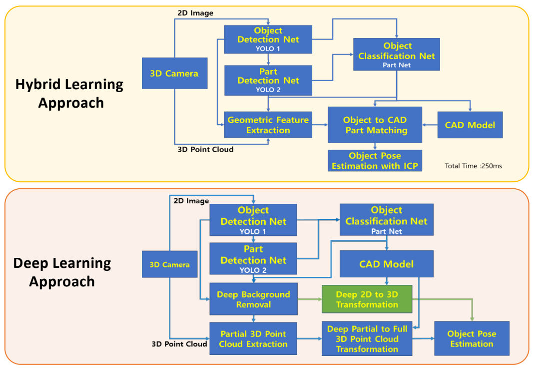
Industrial Bin Picking

30

AI-Robot Skills: Industrial Bin Picking

Industrial Bin Picking

- Machine learning approach to the real-time bin picking of randomly filed 3D industrial parts
  - Hybrid Approach
  - Deep Learning Approach

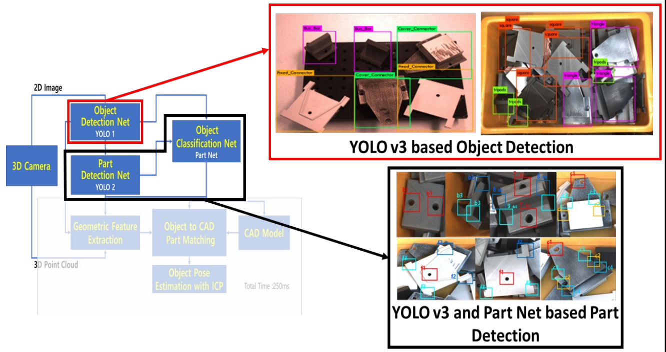


31

AI-Robot Skills: Industrial Bin Picking

Industrial Bin Picking

- Deep Learning Objects and Object Parts Detection



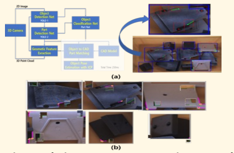
32

AI-Robot Skills: Industrial Bin Picking

Industrial Bin Picking

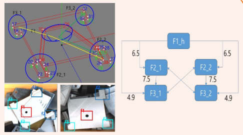
➢ Hybrid Approach

- **Feature Matching and 3D Pose Estimation**
  - ✓ Extracted part features from the cascaded object/part and part feature detector are conveniently used to extract the point and line features.
  - ✓ The point and line features are involved in individual part features that can be assigned with 3D data by incorporating 3D point cloud data from 3D camera.



Point and Line Feature Extraction Based on the Object/Part Features Detected by Object/Part Feature Detector Net

- **Part Matching with CAD Model**
  - ✓ Based on the extracted part features, we can approximated estimate the object pose by matching them with the pre-defined CAD features



Matching between the extracted part features and the pre-defined CAD features for pose estimation

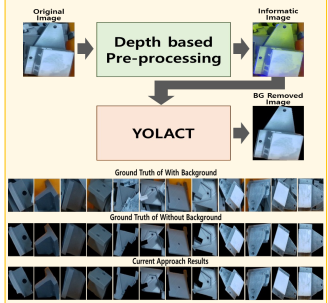
33

AI-Robot Skills: Industrial Bin Picking

Industrial Bin Picking

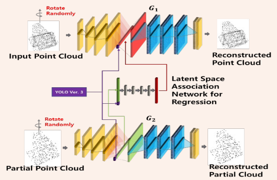
➢ Deep Learning Approach

- **Deep Learning Approach:**
  - ✓ Background Removal of Object/Part Images
  - ✓ Pre-process the Image by Depth based Tone Mapping Algorithm

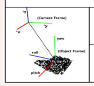


Results of Background Removal

- **Deep Learning Approach:**
  - ✓ Deep Partial-to-Full 3D Point Cloud Transformation



Partial-to-Full Transformation Network based on Bias-Induced Point AE



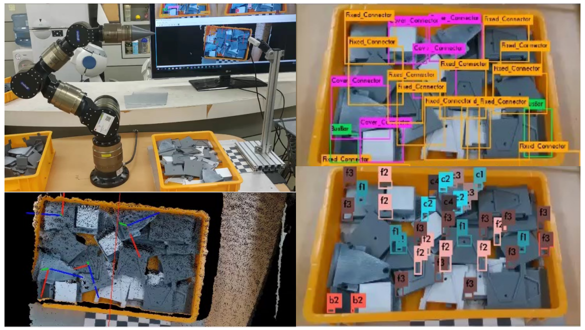
		Average Mean Square Error (MSE)		
	Position Error (cm)	ΔX_0	ΔY_0	ΔZ_0
		0.21	0.20	0.33
	Orientation Error (degree)	ΔRoll	ΔPitch	ΔYaw
		16.82	15.15	21.04

The pose estimation Performance based on Deep Learning approach

34

AI-Robot Skills: Industrial Bin Picking

Industrial Bin Picking - Demo



35

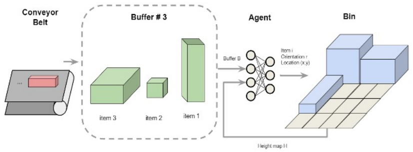
AI-Robot Skills:

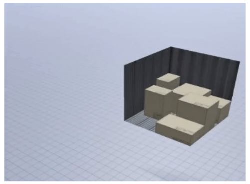
Industrial Bin Packing

36

3D Bin Packing Problem

- **Problem:** recommend an item *i* stored in a buffer, orientation *r* and a discrete loading position *(x, y)* at each time step *t*.
- **Goal:** optimize space utilization.
- **Constraints:** (1) Instability check for safe placement. (2) The cuboid items are constraints in size to half of the bin dimensions.
- **Challenges:** (1) NP-hard combinatorial problem. (2) Incomplete knowledge of the conveyor sequence during inference.





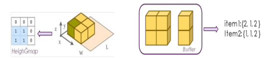
37

Approach 1

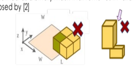
AI-Robot Skills: Industrial Bin Packing

- **A new modality based on item buffering**
- **A novel data augmentation to improve on sampling efficiency**
- **Showing competitive results with an AlphaGo inspired model-based reinforcement learning approach.**

- The **state** is defined by the **bin state** as the 2D heightmap and buffer items dimensions



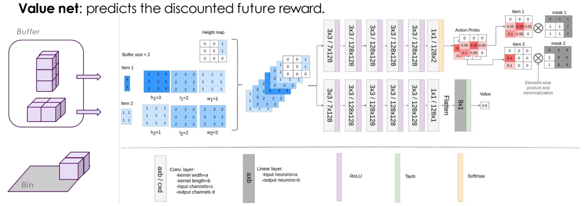
- **Action** is defined as the placement location taking as reference the front-left-bottom of the item (FLB). We generate the probability for each placement for each item and orientation available.
- **Feasibility mask:** is defined as a binary mask that removes invalid actions, due to instabilities or exceeding the bin dimensions. Proposed by [2]



- **Reward**: step-wise reward is equal to $r_i = 10\, V_i/V_B$ where V_i is the items volume to pack and V_B is the bin volume.

The architecture is composed of two subnets:

- **Policy net**: predicts the action probability placements for every location, orientation and item.
- **Value net**: predicts the discounted future reward.



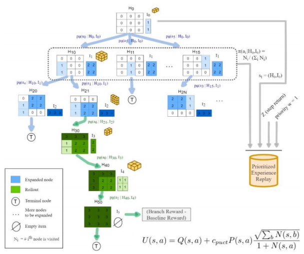
38

Approach 2

AI-Robot Skills: Industrial Bin Packing

Experience Collection

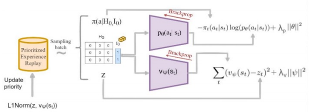
1. Given the heightmap Ht and items I_t at time t.
2. Compute the greedy accumulated score from the policy at time t (Baseline reward).
3. Run **MCTS** for a fixed number of simulations with rollouts to obtain an approximation to the opti mal policy
4. Store the action probs, current state, and ex pected return at a the **priority experience r eplay**.



$$U(s,a) = Q(s,a) + c_{puct} P(s,a) \frac{\sqrt{\sum_b N(s,b)}}{1 + N(s,a)}$$

Update the Policy

1. Sample a batch from the prioritized experience replay
2. Forward the state though the policy and the value nets.
3. Backpropagate the the loss with the proposed objectives.
4. Update the priority of the samples according to for instan ce the value loss.



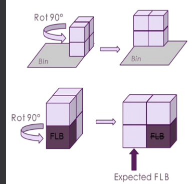
39

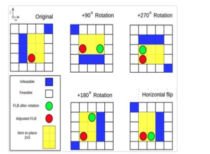
Approach 3

AI-Robot Skills: Industrial Bin Packing

Data Augmentation

- We apply **rotation** and **flip** to both items and bin. Neither operations alternates the expected space utilization.
- **Augmenting** training samples up to **x8**.
- FLB is changed after the operations, as a results we need to re-adjust the action probabilities.

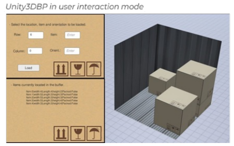




Examples of FLB displacement after applying different actions

- **Unity3DBP:** Unity3D based with 3D visualization and physics engine
- **Py3DBP:** minimalistic python based environment.

Unity3DBP in user interaction mode



40

AI-Robot Skills: Industrial Bin Packing

Results

Empirical Results

Datasets where the bin is 10x10x10, proposed by [2]:

- **k** = # orientations - 1, **b** = Buffer size.
  - **CUT-1:** cutting-stock with items delivered according to the *vertical axis*.
  - **CUT-2:** cutting-stock with items delivered according to the *stacking dependency*.
  - **RS**: random generated items, perfect packing is not guaranteed.

Observations:

- The larger the buffer the more performance
- Our method outperforms other works across multiple datasets.

k=0, b=1	CUT-1	CUT-2	RS
Heuristics [8]	15.15/59.8%	17.3/61.19%	**13.8/54.3%**
Model-free [2]	19.1/73.4%	17.5/66.9%	12.2/50.5%
Ours	**21.3/83.4%**	**18.0/69.9%**	13.1/53.1%
k=1, b=1			
Heuristics [8]	15.99/61.5%	17.62/62.5%	13.82/56.9%
Model-free [2]	19.4/76.2%	18.1/70.2%	15.2/62.1%
Ours	**22.1/85.6%**	**20.2/73.9%**	**15.7/64.2%**

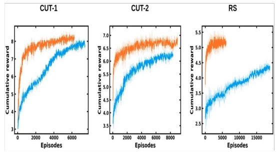
Single-item selection (b=1)

	CUT-1	CUT-2	RS
b=1	21.3/83.4%	18.0/69.9%	13.1/53.1%
b=2	22.0/84.0%	20.2/71.5%	14.6/57.6%
b=3	**22.5/85.7%**	**21.8/77.1%**	**15.8/62.1%**

Multi-item selection (b>1)

Effect of Data Augmentation

- Faster convergence!
- Better performance.



With augmentation W
ithout augmentation

41

Thank You!

42