General

ID[1]			
Use case name	Robotic prehension of objects		
Context	Other (please specify) Robotics		
Application domain	Embedded systems		
Status	PoC		
Contributor	Name	Affiliation	Contact
Scope[2]	Outputting end effector velocity & rotation vector in response to view from RGB-D camera located on robot wrist		
Objective(s)	Use reinforcement learning to train the robot to grasp misc. objects in simulation and transfer this learning to real-life robots.		
Narrative	Short description (not more than 150 words)	It may be difficult and time-consuming for clients of assistive robotic arms to control them with the fine degree required for grasping household objects (such as in the context of having a meal). In order to improve their quality of life, we propose a method by which users can select the bounding box around the object they wish grasped, and the robot performs the grasping action. We use methods from reinforcement learning to train first in simulation, in order to reduce total training time and potential robot breakage, and then transfer this learning to real-life.	
	Complete description	It can be very difficult and time-consuming for users to perform fine movements with a robot arm, like grasping various household objects. To mitigate this problem, attempts are made to grant users the ability to control the arm at a higher level of abstraction; thus, rather than specifying each translation and rotation of the arm, we would like them to be able to select an object to grasp, and have the arm grasp it automatically. This requires some degree of computer vision, to be able to detect objects in the robot's field of view (a camera will be affixed to its wrist). With that achieved, we will be able to focus on grasping an object selected from the detections. Current literature on robotic grasping One might be tempted to start from a heuristic, geometric approach. That is, to use a set of pre-established rules for picking up objects -- for example, executing pincer grasps from the top along the thinnest dimension of the object that is not too narrow to be grasped. Such approaches work reasonably well in conditions that match the restrictive assumptions on which the rules are built, but fail when encountering even small deviations from those conditions (for example, they do not adapt well to clutter). Attempting to list and plan a proper response to all such failure cases heuristically would be an exercise in futility. In contrast, approaches based on machine learning can generalize to unforeseen or novel situations, and, as in the case of object detection, generally perform better than heuristic solutions. Machine learning-based approaches to grasping and object manipulation vary widely. At the simplest level, we can predict the likelihood of grasp success based on an image patch of an object and a given angle of approach. Robot control, in such cases, is beyond	

		the scope of the machine learning model. However, methods can scale up to end-to-end systems which learn to control the robot at the level of its joint actuators in response to a visual stimulus consisting of a bird’s eye view of the arm and several objects placed in a bin.		
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Success rate in simulation	Grasp success rates on both objects seen during training, and new objects, in simulation.	Improve accuracy and generalization.
	2	Success rate in real life	Grasp success rates on both objects seen during training, and new objects, in real life.	Improve accuracy and generalization.
AI features	Task(s)	Planning		
	Method(s)[3]	Reinforcement learning, deep learning		
	Hardware[4]	Depth camera, RGB camera, GPU, actuators, gripper		
	Terms and concepts used[5]	Reinforcement learning, Deep learning, point cloud, depth, scene completion, grasping, transfer learning		
Challenges and issues	Challenges: The camera cannot have a bird's eye view and will instead move with the robot. Sparse rewards may complicate learning. Environment may be cluttered, occlusions of target might occur, objects may move around Issues: For safety reasons, speed and force of robot need to be limited in assistive environment to avoid harm. Human intervention can happen at any time.			
Societal concerns	Prevent arm to people and animals near robot when it is performing a grasping task			

Data (optional)

Data characteristics	
Description	
Source[6]	
Type[7]	
Volume (size)	
Velocity (e.g. real time)[8]	
Variety (multiple datasets)[9]	
Variability (rate of change)[10]	
Quality[11]	

Process scenario (optional)

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition[12]	Post-condition[13]

Training (optional)

Scenario name	Training				
Step No.	Event[14]	Name of process/Activity[15]	Primary actor	Description of process/activity	Requirement

Specification of training data[16]	

Evaluation (optional)

Scenario name	Evaluation				
Step No.	Event[17]	Name of process/Activity[18]	Primary actor	Description of process/activity	Requirement

Input of evaluation[19]	
Output of evaluation[20]	

Execution (optional)

Scenario name	Execution				
Step No.	Event[21]	Name of process/Activity[22]	Primary actor	Description of process/activity	Requirement

Input of Execution[23]	
Output of Execution[24]	

Retraining (optional)

Scenario name	Retraining				
Step No.	Event[25]	Name of process/Activity[26]	Primary actor	Description of process/activity	Requirement

Specification of retraining data[27]	

References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link

Footnote

[1] Leave this cell blank.

[2] The scope defines the limits of the use case.

[3] AI method(s)/framework(s) used.

[4] Hardware system used.

[5] Terms and concepts listed here can be used to extend the work of WG 1 (AWI 22989 and AWI 23053) as necessary.

[6] Origin of data, which could be from instruments, IoT, web, surveys, commercial activity, or from simulations.

[7] Structured/unstructured Images, voices, text, gene sequences, and numerical. Composite: time-series, graph-structured

[8] The rate of flow at which the data is created, stored, analysed, or visualized.

[9] Data from a number of domains and a number of data types. The wider range of data formats, logical models, timescales, and semantics complicates the integration of the variety of data.

[10] Changes in data rate, format/structure, semantics, and/or quality.

[11] Completeness and accuracy of the data with respect to semantic content as well as syntactical of the data (such as presence of missing fields or incorrect values)

[12] Describe which condition(s) should have been met before this scenario happens.

[13] Describe which condition(s) should prevail after this scenario happens. The post-condition may also define "success" or "failure" conditions.

[14] The event that triggers the step. This might be completion of the previous event.

[15] Action verbs should be used when naming activity.

[16] Training data can be further specified.

[17] The event that triggers the step. This might be completion of the previous event.

[18] Action verbs should be used when naming activity.

[19] Specify input of evaluation.

[20] Specify output of evaluation.

[21] The event that triggers the step. This might be completion of the previous event.

[22] Action verbs should be used when naming activity.

[23] Specify input of evaluation.

[24] Specify output of evaluation.

[25] The event that triggers the step. This might be completion of the previous event.

[26] Action verbs should be used when naming activity.

[27] Retraining data can be further specified.