This article has been accepted for publication in IEEE Transactions on Cognitive and Developmental Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2023.3284070

Continual Learning for Anthropomorphic Hand Grasping

Wanyi Li, *Member, IEEE*, Wei Wei, *Student Member, IEEE*, Peng Wang, *Member, IEEE*

***Abstract*—It is important and challenging to learn to grasp different objects with anthropomorphic robotic hands continually and incrementally. However, most current works do not have this property: They learn grasp planners using large pre-prepared datasets, do not generalize well to new objects, and are difficult to improve continually. Besides, existing continual leaning works rarely target at anthropomorphic hand grasping, and usually deal with short streams of experiences. Because of the intrinsic long stream nature of anthropomorphic hand grasping, it is hard to utilize off-the-shelf continual learning methods for it. In this paper, we propose to introduce continual machine learning into anthropomorphic hand grasping and design the Continual Learning Framework of Anthropomorphic Grasping (CLFAG framework). It includes three modules: Data Producer, Grasp Experiences, and Continual Learning Algorithm A_{CL}, thus makes the continual learning of anthropomorphic grasping possible. To overcome the catastrophic forgetting problem in long streams of grasping experiences, we propose a continual learning algorithm based on importance-based regularization and diversity-aware replay within the CLFAG framework. Furthermore, we construct a dataset for continual learning of anthropomorphic grasping. Experiments on constructed dataset and in simulation demonstrate the effectiveness and superiority of the proposed approach.**

***Index Terms*—Continual learning, anthropomorphic hands, grasping.**

I. Introduction

GRASPING is one of the most fundamental skills for robots to interact with objects, because robots usually need to grasp an object in most of manipulation tasks. Despite great progress has been made in grasping with parallel-jaw grippers, parallel grippers can only perform simple object interactions. Multi-fingered anthropomorphic hands are able to perform human-like grasping across various categories of objects. In ever-changing and unstructured environment, robots need to learn new knowledge over time as it is not possible to pre-program everything in advance. The capability to learn new knowledge and abilities over time without forgetting the previously learnt knowledge is referred to as continual learning (CL). Endowing anthropomorphic robotic hand with the ability to learn to grasp different objects continually over time, could have enormous societal impact and have immense potential applications in unstructured environment. For example, to provide assistance in household of disabled or elder people, and to resort and package various goods in factories.

Previous works on anthropomorphic robotic hand grasp generation can be divided into two groups: analysis-based approaches and learning-based approaches. Analysis-based approaches rely on the assumption of a fully-observable environment and the simplification of the physical models [6], that could scarcely be achieved in real-world scenarios. Recent years, with the progress of machine learning, especially deep learning, learning-based approaches for anthropomorphic robotic grasping have made significant breakthroughs and are gradually dominating the community. By utilizing supervised learning or reinforcement learning paradigms, these learning-based approaches train the grasping policy with large amounts of annotated data, of which grasp annotations are collected by humans [52], with simulation [49] or physical robot tests [13]. With a large amount of training data, the performance of learning-based approaches substantially outperforms the classical analysis-based methods. Although current learning-based grasping approaches can learn surprising grasp ability, but they do not have continual learning capability. They usually base on current deep neural network learning models and require large batch of annotated samples to train grasp model. Catastrophic forgetting is a typical issue of current deep neural network learning models [37]. This phenomenon typically leads to a significant performance decrease when neural network models are trained on sequential experiences or tasks with samples becoming progressively available over time. Consequently, these learning-based grasping models do not generalize well to new objects, and are difficult to improve continually.

Continual learning aims at alleviating catastrophic forgetting in learning process. Existing continual learning works include regularization-based, rehearsal-based, and architectural methods [37]. They mainly focus on image classification tasks. Some works deal with continual learning of other tasks such as robot perception [27], reinforcement learning [43] and language processing [20]. However, introducing con-

Manuscript received Month, Day, 2022; This work was supported by the National Key Research and Development Program of China (2018AAA0103003), the National Natural Science Foundation of China (61771471, 91748131 and 62006229), the Strategic Priority Research Program of Chinese Academy of Science under Grant XDB32050106, and in part by the InnoHK Project. (Corresponding author: Peng Wang).

Wanyi Li is with State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail:wanyi.li@ia.ac.cn).

Wei Wei is with State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: wei.wei2018@ia.ac.cn)

Peng Wang is with State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, with the CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China, and also with the Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science and Innovation, Chinese Academy of Sciences, Hong Kong 999077, Hong Kong (e-mail: peng_wang@ia.ac.cn).

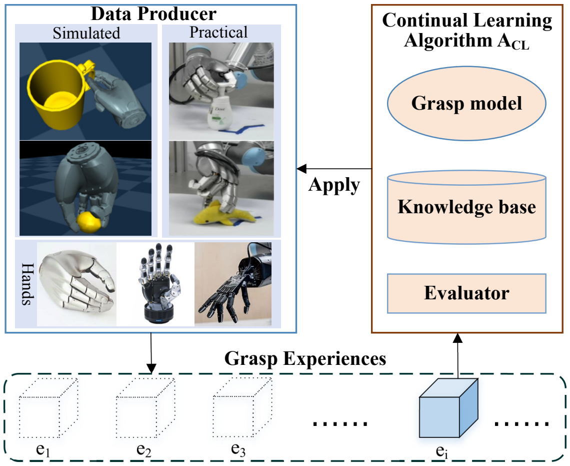


Fig. 1. Continual learning framework of anthropomorphic grasping.

tinual learning into anthropomorphic grasping has not been investigated up to now. Moreover, existing CL works usually deal with short streams of tasks or experiences. In contrast, anthropomorphic grasping is with longer streams of grasping experiences, in which the grasps of a great many objects have to be learnt. It is difficult to utilize off-the-shelf continual learning methods for anthropomorphic hand grasping.

In this paper, we propose to introduce continual machine learning into anthropomorphic hand grasping and design the Continual Learning Framework of Anthropomorphic Grasping (CLFAG framework), as shown in Fig.1. Three modules are included in the designed CLFAG framework, i.e., Data Producer, Grasp Experiences, and Continual Learning Algorithm A_{CL}, which enable anthropomorphic robotic hands to learn to grasp different objects continually and incrementally over time. To overcome catastrophic forgetting in long grasping stream, we propose a continual learning algorithm within the CLFAG framework. The proposed algorithm consists of two major components: importance-based regularization and diversity-aware replay. Importance-based regularization regularizes the update of model parameters according to their importance for previous experiences, while diversity-aware replay maximizes the diversity of replayed samples. The two components interacts with each other by mixing buffered samples and incoming samples of current task, and updating grasp model with a regularized loss function.

Besides, we construct a dataset for continual learning of anthropomorphic robotic hand grasping. The dataset includes more than one million grasp samples using an anthropomorphic robot hand, i.e., DLR/HIT Hand II [22], on 270+ objects. With regard to continual grasp learning setting, the grasping sample set for each object is taken as a new separate experience to learn continually. We evaluate the proposed approach and compare it with several mainstream continual learning approaches on our constructed dataset and in simulation. The results on dataset demonstrated that our methods are more capable of alleviating catastrophic forgetting in long stream of grasping experiences, and are with lower loss. Furthermore, the results on grasping seen objects from YCB and unseen objects from EGAD! in simulator MuJoCo shown that, our method achieves significant improvement compared to baselines in terms of grasp success rate and hand-object penetration.

The main contributions of this paper are summarized as follows:

1) We propose the Continual Learning Framework of Anthropomorphic Grasping (CLFAG framework), which introduces continual machine learning into anthropomorphic hand grasping, and enables anthropomorphic robotic hands to learn to grasp different objects continually and incrementally over time. To the best of our knowledge, this is the first work to deal with continual learning of anthropomorphic hand grasping.
2) We propose a hybrid algorithm integrating importance-based regularization with diversity-aware replay for continual learning of anthropomorphic grasping, which can prevent forgetting and preserve grasp knowledge well over long stream of grasping experiences.
3) We construct a dataset for continual learning of anthropomorphic grasping. We evaluate the proposed approach and compare it with other CL methods on our dataset and in simulation. Experimental results demonstrate the effectiveness and superiority of the proposed approach.

We organize the remainder of the paper as follows. In Section II, we discuss the related works. In Section III, we describe the proposed continual learning framework of anthropomorphic grasping. In Section IV, we explain the detailed methodology of our continual learning algorithm. In Section V, we report and discuss the experimental results. Finally, we draw conclusions and suggest possible future direction in Section VI.

II. Related Work

In this section we discuss the most relevant works including learning-based anthropomorphic robotic hand grasp generation and continual learning. More thorough reviews on continual leaning can be found in [37][27][43], and reviews on grasp learning in [7][11][24][38].

A. Learning-based anthropomorphic hand grasping

Learning-based anthropomorphic grasping approaches are typically categorized into supervised learning and reinforcement learning. The grasp annotations for supervised learning are obtained either from humans [52], simulation [49], or real robot tests [13], while the supervisory signals for reinforcement learning are collected with interactive trial and error.

Supervised learning for anthropomorphic grasping. According to whether the grasp configuration is the input or the output of the learnt model, supervised anthropomorphic grasp learning can be divided into discriminative or generative. Discriminative methods sample grasp candidates and rank them use a network. For example, [2][25] train evaluation models to rank grasps generated by a grasp sampler. [33] learns a prior over grasp configurations as a mixture-density network (MDN) to sample grasp candidates, and learns a grasp success prediction model with a voxel-based 3D convolutional neural network (CNN) to rank grasps. Generative methods generate grasp configuration directly. For instance, [31][30] utilize deep neural networks to learn a one-to-one mapping from voxels or depth image of object to high-DOF grasp configuration. [49] presents deep variational grasp generation (DVGG) to generate high-quality diverse grasp configurations from single-view observation, which is based on the Conditional Variational Auto-Encoder (CVAE).

Reinforcement learning for anthropomorphic grasping. Deep reinforcement learning learns control policies by trial and error, and has shown promising results. Taking sensory signals as inputs, dexterous grasping may be performed. [50] proposes a Generative Attention Learning (GenerAL) framework for multi-fingered grasping in clutter, in which a policy gradient formulation and a learnt attention mechanism are utilized. [13] proposes demonstration augmented reinforcement learning for grasping with anthropomorphic hands. [35] proposes to learn dexterous grasping by embedding an object-centric visual affordances within a deep reinforcement learning loop. [29] proposes a mutifingered grasping method based on multimodal reinforcement learning, in which fingertip tactile sensing, joint torques and proprioception are combined to train the multimodal agent in simulation.

Although great progress has been made in learning-based anthropomorphic grasping, existing learning-based approaches generally require large amounts of pre-prepared data to train grasp planners, do not generalize well to unseen objects, and are difficult to improve continually after deployment. This work focuses on continual Learning of anthropomorphic grasping, and aims to endow the anthropomorphic robotic grasping models with the ability to learn grasping different objects continually and incrementally over time.

B. Continual learning

Humans can learn new knowledge and ability throughout lifetimes continually without catastrophically forgetting what they have learnt previously. The ability is referred to as continual learning (CL). However, conventional machine learning and neural network models mostly depend upon fixed datasets and stationary environments, struggle to learn from non-stationary data distributions over time. CL remains a great challenge for conventional learning algorithms, since continual acquisition of gradually available information from non-stationary data distributions generally causes catastrophic forgetting or interference. CL focuses on alleviating catastrophic forgetting while accommodating new information effectively. CL algorithms can be categorized into three groups: regularization-based, rehearsal-based, and architectural methods.

Regularization-based CL methods. Regularization approaches impose constraints on the update of network parameters to alleviate catastrophic forgetting. Typical examples include elastic weight consolidation (EWC) [23], Synaptic Intelligence (SI) [51], and Knowledge Distillation (KD) [19][28]. EWC [23] preserves knowledge of previous tasks by selectively slowing down updating on the weights important to those tasks, the importance of parameters is measured via the diagonal of the Fisher Information Matrix. SI [51] estimates importance of individual synapses in an online fashion over the entire learning trajectory, and penalizes changes to the most important synapses, thus learns new tasks with minimal forgetting. KD [19][28] preserves a model's functionality for old tasks by encouraging the output of previous task layers to be consistent even when learning a new task. The main drawback to regularization-based approaches is that it is hard to prevent forgetting when learning from long task sequences [16][27]. Long task sequence means that the model is presented with many tasks or many experiences.

Replay-based CL methods. To mitigate forgetting, replay strategy stores a representation of previous data to combine with new data while update the neural network. Two ways have been used in context of continual learning: partial replay and generative replay. For partial replay, all or a subset of previously learnt input is stored in a replay buffer. For example, several successful models store a subset of raw input of previous tasks in a replay buffer [9][4][10][45]. Some methods that store intermediate representation for replay have been proposed also [18][15][8]. Different from storing previous examples explicitly, generative replay approaches train a generative model such as a generative adversarial network (GAN) [14] or an auto-encoder from previously learnt data, and use it to generate samples. For instance, [44][1] generate veridical inputs and [47][26] generate mid-level CNN feature representations. The good news for replay-based methods is that they are effective in continual learning, while the bad news including the performance is limited by available compute and memory resource, and they tend to overfit on the replay memory thus impairing generalization.

Architectural CL methods. The architectural methods expand model architectures flexibly to accommodate new information while keeping part of the architectures trained for preceding tasks. In architecture method, different parts of model parameters are assigned to different tasks. According to whether the model structure is expanded, architectural methods can be categorized into two groups, namely, fixed networks and dynamic networks. The fixed network methods permit only inner adjustments such as changes in the weights and activations. The representative methods of fixed networks include PackNet [34] and PathNet [12]. PackNet [34] uses iterative pruning and network re-training to add multiple tasks to a single network. PathNet utilizes evolutionary strategies to select pathways that decide the network parameters to be retrained or updated. PathNet [12] fixes the parameters along a path learnt on the previous task and evolves a new group of paths for the new task. Dynamic networks learn new task by adding new modules to the model while the previous learnt parameters are kept unchanged. For example, progressive neural networks (PNN) [40] keeps a group of pre-trained model for previous learnt tasks, from which PNN extracts good features by learning lateral connection for model of the new task. RPSnet [39] employs a random path selection algorithm to select optimal paths for the new tasks while encouraging parameter sharing and reuse. Architectural methods may be quite effective in terms of performance metrics and reducing forgetting. However, they are usually difficult to perform efficient knowledge transfer and parameter sharing, and involve uncontrolled growth in the parameter space.

Existing works of continual learning focus on alleviating catastrophic forgetting and are mainly oriented towards image classification tasks. Introducing continual learning into anthropomorphic grasping has not been investigated up to now also. Moreover, image classification tasks in existing CL works are usually with short task sequence. In contrast, anthropomorphic grasping are with long sequence. Because of these factors, instead of utilizing contemporary CL methods mentioned above directly, we propose to introduce continual machine learning into anthropomorphic hand grasping, design the Continual Learning Framework of Anthropomorphic Grasping (CLFAG framework), and explore methodologies from the literature that are suitable for continual learning of anthropomorphic grasping [51][9], adapting and integrating methods according to the long stream challenge of grasping as continual learning algorithm of anthropomorphic grasping.

III. Continual Learning Framework of Anthropomorphic Grasping

To enable anthropomorphic robotic hands to learn to grasp different objects continually and incrementally over time, we design the Continual Learning Framework of Anthropomorphic Grasping (CLFAG framework), which is shown in Fig.1. The CLFAG framework provides the conceptual foundation for continual learning of anthropomorphic grasping. There are three modules in the framework: Data Producer, Grasp Experiences, and Continual Learning Algorithm A_{CL}. Data Producer generates a stream of anthropomorphic grasping experiences e_i. Grasping Experiences e_i are sequentially accessible by the Continual Learning Algorithm A_{CL} with its internal grasp model and knowledge base. Directly interacting with the algorithm A_{CL}, the evaluator computes performance metrics p_i. Detail of the CLFAG framework is described below.

In CLFAG framework, data is modeled as a sequential non-iid learning experiences, i.e., anthropomorphic grasping experiences:

$$\mathbf{E} = (e_1, e_2 \ldots, e_n) \tag{1}$$

Anthropomorphic grasping experiences are generated by the Data Producer in simulations or practical systems. A learning experience e_i consists of a set of grasp samples of an object which can be used to update the grasp model. A grasp sample is represented as $\langle \mathcal{P}, \boldsymbol{g} \rangle$, where $\mathcal{P}$ is the observation of an isolated object and $\boldsymbol{g}$ is a grasp. Possible options of $\mathcal{P}$ include RGB image, point cloud and RGB-D image. The model M takes the observation $\mathcal{P}$ as input and predicts high quality grasps. Each grasp $\boldsymbol{g}$ is expressed by a hand wrist pose $\boldsymbol{p}$ and a hand joint configuration $\boldsymbol{\theta}$, i.e. $\boldsymbol{g} = \{\boldsymbol{p}, \boldsymbol{\theta}\}$. The hand wrist pose $\boldsymbol{p}$ is given in special Euclidean group $\mathcal{SE}(3)$, consisting of the translation $\mathbf{t} = [t_x, t_y, t_z]$ and orientation quaternion $\mathbf{q} = [q_w, q_x, q_y, q_z]$. The hand joint configuration $\boldsymbol{\theta}$ is denoted by the actual degree of freedom of the anthropomorphic robot hand, for example, $\boldsymbol{\theta} \in \mathbb{R}^{20}$ for anthropomorphic robot hand DLR/HIT Hand II. The proposed CLFAG framework supports different anthropomorphic robot hands such as Shadow Dexterous Hand [42], Schunk five-finger hand [41], and DLR/HIT Hand II [22]. This paper uses DLR/HIT Hand II as a case study for implementation of the CLFAG framework.

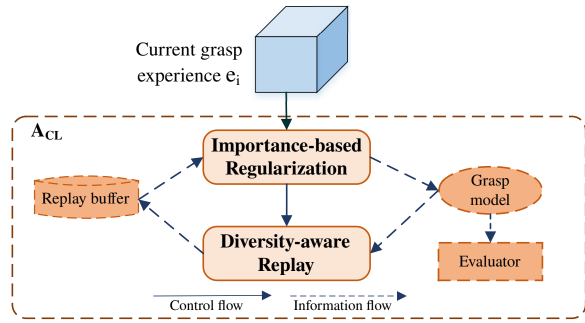


Fig. 2. Schematic view of the proposed hybrid algorithm for continual learning of anthropomorphic grasping.

During training, a continual learning algorithm A_{CL} processes grasp experiences sequentially and uses them to update the internal grasp model and the knowledge base. The knowledge base is usually represented as specific data structures, such as replay buffer and neural network's trained weights. In CLFAG framework, each CL algorithm has a training mode and an evaluation mode. Training mode is used to update the model, and evaluation mode may be used to process streams of experiences for test purpose.

The continual learning framework of anthropomorphic grasping can be formalized as follows.

Definition. (Continual Learning Framework of Anthropomorphic Grasping). A continual learning algorithm A_{CL} is expected to update its internal state, e.g. its internal grasp model M and knowledge base, based on a non-stationary sequentially accessible stream of anthropomorphic grasping experiences $(e_1, \ldots, e_n)$. The grasping experiences can be obtained from simulations or practical systems. The objective of A_{CL} is to improve its performance on a set of grasp metrics $(p_1, \ldots, p_m)$ as evaluated on a test stream of experiences $(e_1^t, \ldots, e_n^t)$.

IV. The Proposed Hybrid Continual Learning Algorithm

We first describe our algorithm at a high level here. It consists of two major components: importance-based regularization and diversity-aware replay. The schematic view of the proposed algorithm is given in Fig.2. Importance-based regularization preserves previous knowledge in model parameter space by updating model parameters according to their importance, while diversity-aware replay keeps knowledge in sample space by maximizing the diversity of samples in the replay buffer. Importance-based regularization interacts with diversity-aware replay by mixing buffered samples and samples of current experience to update the grasp model. We provide details of two major components in the following subsections.

A. Importance-based regularization

To overcome catastrophic forgetting in parameter space, the importance-based regularization component identifies important parameters for previous tasks and slows down their update. Our importance-based regularization relies on synaptic intelligence [51], a continual learning method that computes the importance of each parameter online and penalizes changes to important synapses. We mix the buffered samples and samples of current experience, and perform update to the grasp model by adding an importance-based regularization term to the grasp loss with respect to the mixed samples.

More specifically, for each incoming mini-batch $\boldsymbol{B}_n$ drawn from current grasp experience e_i, the algorithm retrieves and augments another mini-batch $\boldsymbol{B}_m$ from the replay buffer $\mathcal{M}$. The retrieval and the augmentation of B_m is described in diversity-aware replay component, i.e., Section IV-B. We merge $\boldsymbol{B}_n$ and $\boldsymbol{B}_m$. And then we perform the model update by optimizing the following loss with respect to the parameters of the grasp model M on merged sample set $\boldsymbol{B}_n \cup \boldsymbol{B}_m$.

$$\widetilde{\mathcal{L}_i} = \mathcal{L}_i(\boldsymbol{B}_n \cup \boldsymbol{B}_m) + c\sum_k \Omega_k^i(\widetilde{\theta_k} - \theta_k)^2, \tag{2}$$

where $\mathcal{L}_i$ is the loss of grasp model M, which is calculated over the merged sample set $\boldsymbol{B}_n \cup \boldsymbol{B}_m$, the regularization term $c\sum_k \Omega_k^i(\widetilde{\theta_k} - \theta_k)^2$ regularizes changes to parameters, of which, c is a strength parameter which trades off old tasks against the new one, k labels each parameter of grasp model, Ω_k^i is the per-parameter importance which is maintained in an online manner during training and is defined as:

$$\Omega_k^i = \sum_{j<i} \frac{\omega_k^j}{(\Delta\theta_k^j)^2 + \xi}. \tag{3}$$

Note that ω_k^j is the parameter specific contribution to changes in the total loss over the entire trajectory of training, detailed computation process can be found in [51]. $\Delta\theta_k^j = \theta_k^j - \theta_k^{j-1}$ is the task-specific parameter distance, the term $(\Delta\theta_k^j)^2$ ensures that the regularization term carries the same units as the loss L, the damping parameter ξ avoids division by zero. Besides, for the strength parameter, $c = 1$ indicates an equal weighting of old and new tasks if ω_k^j is evaluated precisely. However, because the existence of the noise in the evaluation of ω_k^j, c is typically set smaller than one to compensate. Due to the long stream peculiarity of the continual grasp learning, and the kept diverse samples of previous tasks, we argue to use more smaller c to consolidate most important information of old tasks in model parameters only, and to alleviate the saturation-prone property of regularization methods in the long steam.

B. Diversity-aware Replay

We alleviate catastrophic forgetting in sample space with a diversity-aware replay component. Our diversity-aware replay component consists of two steps, namely, memory update with reservoir sampling and memory retrieval using 3D augmentation enhanced random sampling. Below we provide the details of the two steps.

Memory update with reservoir sampling. We adopt reservoir sampling [48] in the memory update step. Reservoir sampling selects random samples from the input stream $\mathcal{S}$ of unknown length. Samples in input stream have same probability of being selected and stored in the replay buffer $\mathcal{M}$. If $|\mathcal{S}|$ is the number of samples observed so far and $|\mathcal{M}|$ is the size of the reservoir (replay buffer), reservoir sampling selects each samples with a $\frac{|\mathcal{M}|}{|\mathcal{S}|}$ probability.

Memory retrieval using 3D augmentation enhanced random sampling. It is argued that the diversity of samples is very important in replay-based continual learning method [3][5]. Given the observed samples $\boldsymbol{B}_n$, we sample a mini-batch $\boldsymbol{B}_m$ from the replay buffer $\mathcal{M}$. To enhance the diversity of retrieved samples, we apply 3D data augmentation on $\boldsymbol{B}_m$. The 3D data augmentation includes jitter, dropout and rotation. Jitter operation adds a clipped Gaussian noise with zero mean and standard deviation σ to the position of each point. Noises larger than t_{clip} threshold are clipped to t_{clip}. The Jitter operation is defined in Eq.(4) and Eq.(5):

$$Jitter(\boldsymbol{B}_m) = \langle \mathcal{P}_m + \boldsymbol{z}, \boldsymbol{g}_m \rangle, \boldsymbol{z} \in \mathbb{R}^{bsize \times n \times 3}, \tag{4}$$

whereby $\mathcal{P}_m$ denotes the observed pointclouds for objects, $\boldsymbol{z}$ is the noise, grasp array $\boldsymbol{g}_m$ remains unchanged, $\boldsymbol{z}$ has same shape $bsize \times n \times 3$ with $\mathcal{P}_m$ of which $bsize$ is the number of samples in $\boldsymbol{B}_m$, $\boldsymbol{z}$ is from a clipped normal distribution and defined as:

$$\boldsymbol{z} = \min(t_{clip}, \boldsymbol{n}), \tag{5}$$

where $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\sigma}^2)$, t_{clip} is the threshold for clip. Dropout augmentation throw away points randomly with max ratio r_{max}. Rotation augmentation randomly rotates the object $\mathcal{P}_m$ and grasp pose $\boldsymbol{p}$ along three axes, and is formulated as follows:

$$\begin{aligned} Rot(\mathcal{P}_m) &= \mathcal{P}_m\boldsymbol{R} \\ Rot(\boldsymbol{g}_m) &= \{\boldsymbol{p}\boldsymbol{R}, \boldsymbol{\theta}\} \\ Rot(\boldsymbol{B}_m) &= \langle Rot(\mathcal{P}_m), Rot(\boldsymbol{g}_m) \rangle, \end{aligned} \tag{6}$$

where $\boldsymbol{R} = \boldsymbol{R}_z(\gamma)\boldsymbol{R}_y(\beta)\boldsymbol{R}_x(\alpha)$ is a random rotation matrix, α, β, γ are the rotated angles along X, Y, Z axis respectively, and are drawn from uniform distribution $U(0, 2\pi)$. An example of 3D augmentation is shown in Fig.3.

V. Experiments

In this section, we evaluate our approach on both our constructed dataset and in simulation against main stream continual learning methods. Firstly we describe the experimental setup including dataset, compared methods, evaluation metrics and implementation details. Then we evaluate the continual learning capability of the proposed approach on dataset. Finally, we compare the grasp performance in simulation.

A. Experimental setup

Dataset. To evaluate our proposed framework and methods for continual learning of anthropomorphic grasping, we construct a sequential anthropomorphic grasping dataset based on our previous work [49]. The used anthropomorphic robotic hand is DLR/HIT Hand II. There are more than one million grasp samples on 300 objects. We firstly remove those objects with few effective grasps, as a result, 278 objects are preserved. And then, we build a continual grasp learning setting, the data in the setting is modeled as an ordered sequence composed of 278 non-iid learning experiences, a learning experience is a

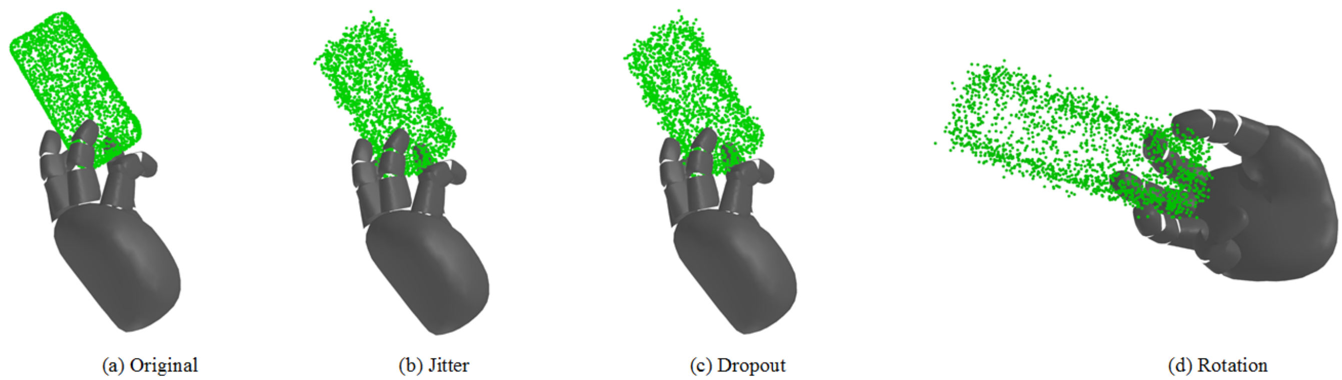


Fig. 3. A example of 3D augmentation.

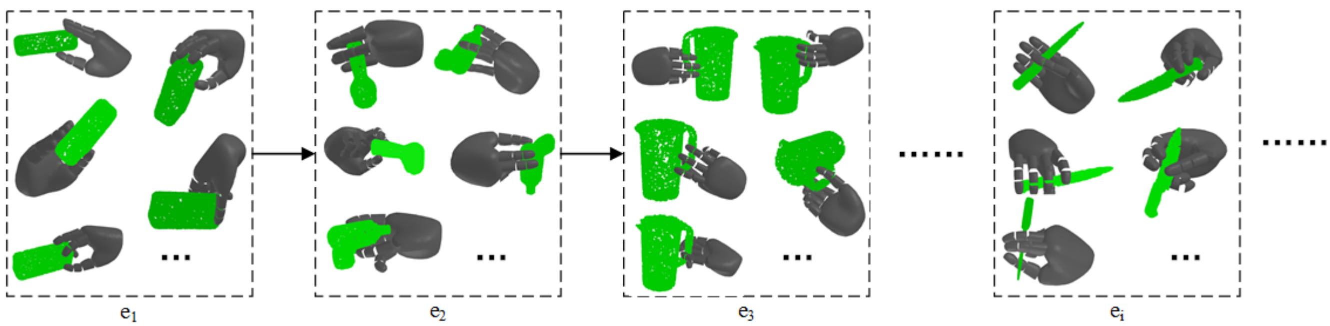


Fig. 4. Grasp samples of each object in dataset as an experience.

set of grasp samples from an individual object, as shown in Fig.4. The complete 3D point cloud of each object is taken as observation $\mathcal{P}$. For each experience, we split the grasp samples into training set, validation set and test set at the ratio of 6:2:2. Training set and validation set in the sequence are used to train the grasping models continually, while the test set is used to test trained models.

Compared methods. We compare the proposed approach with several mainstream continual learning approaches including elastic weight consolidation (EWC) [23], synaptic intelligence (SI) [51] and experience replay (ER) [9], and two baselines, namely, Finetune and IId-Offline. EWC [23] and SI [51] overcome forgetting with importance-based regularization. ER is a simple but effective replay-based approach. It involves storing a portion of previous data and mixing them with more recent ones to update the model. ER applies reservoir sampling [48] for memory update and random sampling for memory retrieval. Finetune incrementally fine tunes the model without employing any continual learning strategy. The model can not overcome catastrophic forgetting and is considered as the lower-bound. IId-Offline uses all the samples in the dataset in an offline manner to train the model, and is regarded as the upper-bound.

Evaluation metrics. For experiment on dataset, we use **Average Loss** (L_i) and **Average Forgetting** (F) to evaluate the continual learning capability of compared methods. The Average Loss (L_i) is the averaged loss of grasp model on test sets of learnt experiences $(e_1, e_2 \ldots, e_{i-1})$ so far after the completion of CL at experience e_i. The **Mean Average Loss** (mAL) is mean of Average Loss over all experiences $(e_1, e_2 \ldots, e_n)$, that is defined in Eq.(7). The Average Forgetting (F) is defined upon Average Loss in Eq.(8).

$$mAL = \sum_{i=1}^{n} L_i \tag{7}$$

$$F = \frac{1}{n-1} \sum_{i=2}^{n} \min(L_i - L_{i-1}, 0) \tag{8}$$

We use three quantitative metrics for evaluation in simulation: **Success Rate (SR)**, the **Penetration Depth** and **Penetration Volume** between the hand mesh and the target object. The three used metrics consistent with previous literature [36][17]. The Success Rate (SR) is commonly used in grasping tasks to measure the stability and quality of the generated grasps. The Penetration Depth is calculated as the max intersection distance between the hand vertices and the object mesh. For the Penetration Volume, we voxelize the hand-object meshes with voxel size 0.5 cm, and calculate the intersection volume shared by the two 3D voxels. To calculate the Penetration Depth and the Penetration Volume, we use the implementation of [21]. When the hand collides with the target object, the Penetration Depth is computed as the maximum of the distances from vertices of hand mesh to the object surface.

Implementation details. For grasp model M of which state expects to be updated by continual learning algorithm A_{CL}, we adopt the variational grasp generator which is the core module of DVGG [49] as a case study. We ignore the two auxiliary steps including object point completion

TABLE I
HYPERPARAMETERS FOR THE COMPARED METHODS.

Methods	Hyperparameter grid	Tuned hyperparameter
Finetune	——	——
IID-Offline	——	——
EWC	$\lambda : [0.1, 1, 100]$	$\lambda = 0.1$
SI	$c : [0.1, 0.5, 1]$	$c = 0.5$
ER	——	——
ER-SI(ours)	$c : [0.1, 0.5, 1]$	$c = 0.5$
ER-DA(ours)	——	$\sigma = 0.005, t_{clip} = 0.025, r_{max} = 0.2$
ER-DA-SI(ours)	$c : [0.1, 0.5, 1]$	$\sigma = 0.005, t_{clip} = 0.025, r_{max} = 0.2, c = 0.5$

and iterative grasp refinement for clarity. The CL algorithms including Finetune, EWC, SI, ER, ER-SI, ER-DA and ER-DA-SI are implemented using Avalanche [32]. For IId-Offline, i.e., the variational grasp generator in DVGG, we use the implementation of [49]. For training of compared methods, 150 epochs is used, and learning rate is set to 0.002 at start and divided by 10 when the validation error plateaus. Batch size is 512. We train all models on a RTX-3090 GPU. We present the detailed hyperparameters in Table I.

B. Results on dataset

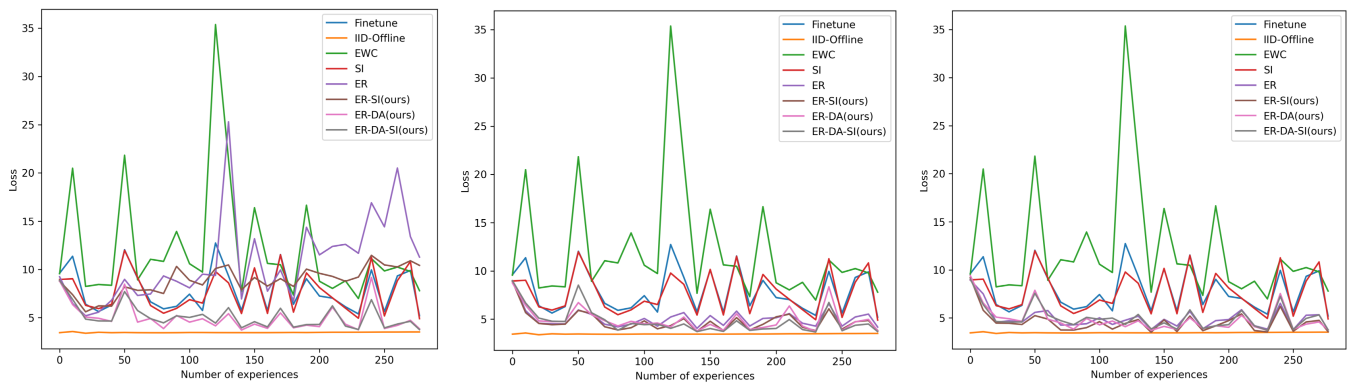
To demonstrate the ability to overcome catastrophic forgetting in long grasping stream of the proposed continual learning algorithm, we compare the the proposed hybrid CL algorithm with other five methods. They are three typical CL methods, namely, EWC, SI and ER, and two baselines, i.e., Finetune and IId-Offline. Our proposed hybrid CL algorithm includes three variants: ER-SI is the integration of experience replay and synaptic intelligence, ER-DA is experience replay with data augmentation, and ER-DA-SI is the integration of experience replay with data augmentation and synaptic intelligence. As test loss of model can reflect the quality of the generated grasps directly, we report and analyze the evolution of test loss along with training. Besides, when all grasp experiences are visited and the grasp model is finally trained, the mean average loss and average forgetting, and the loss on combined test set are also reported and analyzed.

Evolution of test loss along with training. In Fig.5, We show how the loss on test set so far without random rotation evolve along with seeing more tasks, i.e., seeing more objects. In Fig.6, we show evolution of the loss on test set with random rotation. In Fig.5 and Fig.6, the x-axis is the number of experiences. Each experience consists of grasping samples of an object. Acquire more experiences means grasp samples of more objects are visited and learnt.The lower and the smoother the loss is, the better the corresponding continual learning method is. From left to right in Fig.5 and Fig.6, memory size changes with 1K, 5K and 10K. As shown in Fig.5 and Fig.6, the navie Finetune has high loss and oscillates up and down with a large attitude, indicating catastrophic forgetting occurs. It is observed that EWC is even worse than Finetune, the possible causes include: the grasp model is trained from scratch thus the model parameters updated in the early learnt experiences may be not good enough, and after several experiences, too high regularization of EWC makes the grasp model saturated quickly. SI is better than Fineturn, but is still with high loss and large oscillation. ER has high loss and are with large oscillation when the buffer size is small, such as 1K. With the increasing of buffer size, ER performs well gradually. ER is with low loss and small oscillation when big buffer size is used, such as with 5K and 10K replay buffer. By contrast, the variants of our proposed method performs well with lower loss and smaller oscillation. ER-DA-SI achieves best results, which is very close to the IId-Offline, even with only 1K replay buffer. The visualized tendencies of Fig.5 and Fig.6 for alternatives with and without random rotation are similar, indicating that the proposed method is robust to random rotation.

Mean average loss and average forgetting. Table II summarizes Fig.5 and Fig.6 quantitatively by providing the mean average loss and average forgetting. First, observe that compared to EWC, SI and ER, ER-DA-SI with small buffer size 1K shows $\sim 60\%$, $\sim 30\%$ and $\sim 50\%$ relative reduction in mean average loss, $\sim 80\%$, $\sim 60\%$ and $\sim 70\%$ reduction in forgetting, respectively. It is valuable because this indicated that ER-DA-SI was able to work under small memory cost. Under other buffer size conditions, our proposed methods, including ER-SI, ER-DA and ER-DA-SI, consistently out-perform other CL methods with a large margin in terms of mean average loss and forgetting also. Second, similar trend appears both in test set without rotation and with rotation, which indicates the robust of proposed methods to random rotation. Further, compared to IID-Offline, all continual learning methods show over 30% in mean average losses, gap still exists.

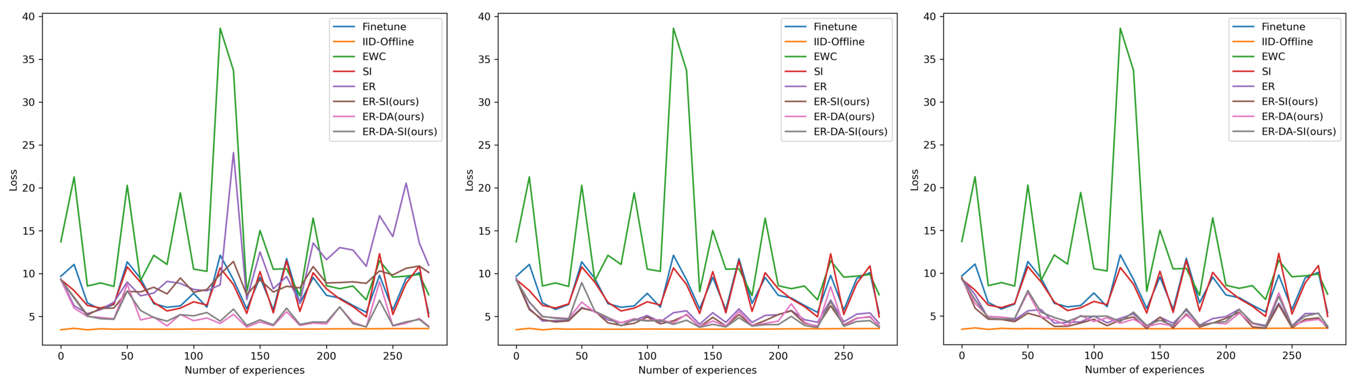
Loss on combined test set of finally trained model. The average loss of finally trained models of all compared methods is shown in Table III. The loss is calculated on combined test set of all experiences. We provide losses of four scenarios, namely, loss on test set without rotation (Loss-test-w/o-rot), loss on test set with rotation (Loss-test-w/-rot), loss on training set without rotation (Loss-train-w/o-rot) and loss on training set with rotation (Loss-train-w/-rot). As shown in Table III, four losses of Fineturn are quite high and are all above 5.2. As expected lower bound, losses of IID-Offline are low and are below 3.6. Corresponding with evolution of test loss along with training in Fig.5 and Fig.6, EWC has high losses around 5.5 which are all larger than those of Fineturn, losses of SI are around 4.9 and are slightly lower than Fineturn's. Losses of ER appear similar tendency with Fig.5 and Fig.6, is high under small buffer size, while is low under big buffer size. For the variants of our proposed method, ER-SI gets lowest loss under 10K buffer size, it is on par with the IId-Offline. ER-DA gets the most significant drop of loss under 1K buffer size, from 11.267 of ER to 3.848. This indicates the data augmentation on retrieved samples enhanced the diversity of replay samples. The full armed version of our proposed method, ER-DA-SI gets mostly lowest loss.

Generally speaking, the proposed CL method performs quantitatively better than the other typical methods on dataset with lower average loss and forgetting. It worth noting that



(a) Buffer Size |M|=1k (b) Buffer Size |M|=5k (c) Buffer Size |M|=10k

Fig. 5. The average loss on test set without random rotation so far measured by the end of each experience (object). (a). 1K buffer size is used for replay related methods. (b). 5K buffer size is used for replay related methods. (c). 10K buffer size is used for replay related methods.



(a) Buffer Size |M|=1k (b) Buffer Size |M|=5k (c) Buffer Size |M|=10k

Fig. 6. The average loss on test set with random rotation so far measured by the end of each experience (object). (a). 1K buffer size is used for replay related methods. (b). 5K buffer size is used for replay related methods. (c). 10K buffer size is used for replay related methods.

TABLE II
MEAN AVERAGE LOSS AND AVERAGE FORGETTING OF COMPARED METHODS. THE BEST RESULTS ARE REPORTED IN BOLD FONT.

Methods	Test-w/o-rot						Test-w/-rot					
	Mean Average Loss $mAL \downarrow$			Forgetting $F \downarrow$			Mean Average Loss $mAL \downarrow$			Forgetting $F \downarrow$		
Finetune	7.889			1.368			7.924			1.284		
IID-Offline	3.470			—			3.520			—		
EWC	12.046			2.816			12.904			2.887		
SI	7.657			1.344			7.686			1.382		
BufferSize$\|\mathcal{M}\|$	1K	5K	10K	1K	5K	10K	1K	5K	10K	1K	5K	10K
ER	10.723	5.094	4.975	1.843	0.406	0.420	10.554	5.147	4.980	1.749	0.416	0.412
ER-SI(ours)	8.853	**4.760**	**4.571**	0.526	**0.380**	**0.401**	8.679	4.828	**4.622**	0.582	**0.394**	**0.409**
ER-DA(ours)	**5.015**	4.978	4.838	0.642	0.504	0.468	**5.015**	5.021	4.850	0.634	0.507	0.463
ER-DA-SI(ours)	5.089	4.799	4.994	**0.509**	0.403	0.518	5.120	**4.826**	5.066	**0.512**	0.421	0.517

there is still a large gap between CL methods and the IID-Offline counterpart, indicating more works for continual learning of anthropomorphic grasping are still expected.

C. Results in simulation

To demonstrate the effectiveness and superiority of the proposed approach on generating anthropomorphic grasps with high stability and quality, we conduct experiments in the physics-based simulator MuJoCo [46]. For comprehensive evaluation, we use 58 objects from YCB dataset (seen) and 48 objects from EGAD! (unseen). For each object, the complete 3D point cloud is provided to the trained grasp model, the grasp model generates 20 grasps randomly. In simulator, we perform grasp with generated grasp configuration for all objects and calculate used metrics. There are four steps in the physical simulation process: 1) Fix the object stationary and initialize the robotic hand with a pre-grasp state, then the hand approaches the object and executes grasping with the generated grasp parameters including hand wrist pose and angles of hand joints until a stable state of the simulator reaches. 2) Then

TABLE III

LOSS ON COMBINED TEST SET OF FINALLY TRAINED MODEL FOR COMPARED METHODS. THE BEST RESULTS ARE REPORTED IN BOLD FONT.

Methods	Loss-test-w/o-rot ↓			Loss-test-w/-rot ↓			Loss-train-w/o-rot ↓			Loss-train-w/-rot ↓		
Finetune	5.299			5.257			5.298			5.256		
IID-Offline	3.519			3.571			3.521			3.572		
EWC	5.501			5.516			5.499			5.514		
SI	4.899			4.938			4.897			4.934		
BufferSize$\|\mathcal{M}\|$	1K	5K	10K	1K	5K	10K	1K	5K	10K	1K	5K	10K
ER	11.267	4.182	3.754	10.982	4.188	3.799	11.283	4.182	3.755	10.991	4.188	3.800
ER-SI(ours)	10.419	3.829	**3.576**	10.122	3.888	**3.586**	10.416	3.832	**3.577**	10.135	3.890	**3.587**
ER-DA(ours)	3.848	3.803	3.771	3.849	3.820	3.799	3.844	3.804	3.772	3.849	3.824	3.798
ER-DA-SI(ours)	**3.763**	**3.710**	3.751	**3.779**	**3.727**	3.759	**3.766**	**3.709**	3.751	**3.782**	**3.730**	3.761

TABLE IV

COMPARISON RESULTS ON GRASPING SEEN OBJECTS FROM YCB IN SIMULATION. THE BEST RESULTS ARE REPORTED IN BOLD FONT.

Methods	Penetration						Success Rate(%) ↑		
	Depth(cm)↓			Volume(cm^3) ↓					
Finetune	0.824			9.142			32.8		
IID-Offline	0.642			7.273			62.0		
EWC	1.014			15.028			36.0		
SI	0.733			8.663			33.6		
BufferSize$\|\mathcal{M}\|$	1K	5K	10K	1K	5K	10K	1K	5K	10K
ER	1.262	0.668	0.608	23.204	8.021	6.340	41.4	53.0	55.4
ER-SI(ours)	1.196	0.664	0.574	22.908	8.108	6.255	48.4	**61.3**	57.1
ER-DA(ours)	0.612	0.584	0.547	6.807	6.499	**6.153**	54.8	55.7	55.5
ER-DA-SI(ours)	**0.552**	**0.532**	**0.527**	**5.711**	**5.855**	6.200	**55.2**	56.1	**58.7**

the gravity is present, fingers keep the grasping force till a stable simulator state reaches or the object falls from the hand. 3) By shaking the hand, the unstable grasps are filtered, and grasps that keep the object in hand are preserved as successful ones. 4) Calculate the metrics including Success Rate (SR), Penetration Depth and Penetration Volume, as mentioned in Sec.V-A.

The comparison results on grasping seen objects from YCB and unseen objects from EGAD! in simulation are shown in Table IV and Table V respectively. As shown in Table IV, Success Rate (SR) on grasping objects from YCB of our proposed ER-DA-SI with 1K, 5K and 10K replay buffer are 55.2%, 56.1% and 58.7%, with increases of 13.8%, 3.1% and 3.3% compared to ER, respectively. Success Rate (SR) of EWC, SI and Fineturn are lower than that of ER even with small buffer size such as 1K. The other two variants of our proposed method, ER-SI and ER-DA, are with higher Success Rate (SR) than ER. All variants of our proposed method are with less penetration than other methods. Similar tendency is shown in Table V also. Overall, the proposed method outperforms other alternatives on grasping object in simulation with higher success rate and lower penetration including depth and volume. Moreover, we observe that ER with 10K buffer, ER-SI, ER-DA, ER-DA-SI outperforms IID-Offline for unseen EGAD! object dataset, perhaps due to the bias from the dominant objects in IID-Offline. Qualitative results shown in Fig.7 and Fig.8 demonstrate that our proposed method is able to generate diverse reasonable grasps.

VI. CONCLUSIONS

In this paper, we propose to introduce continual machine learning into anthropomorphic hand grasping and design the Continual Learning Framework of Anthropomorphic Grasping (CLFAG framework). Within the proposed CLFAG framework, we propose a hybrid algorithm integrating importance-based regularization and diversity-aware replay for continual learning of anthropomorphic grasping. We construct a dataset, and evaluate the proposed approach by comparing it with other CL methods on our dataset and in simulation. The results on dataset demonstrated that our methods are more capable of alleviating catastrophic forgetting in long stream of grasping experiences, and are with lower loss and smaller forgetting. The results on grasping seen objects from YCB and unseen objects from EGAD! in simulator MuJoCo shown that our method achieves significant improvement compared to other alternatives in terms of grasp success rate and penetration.

Starting from this work, some future works are worthwhile considering, for example, neuro-inspired methods for online continual learning of anthropomorphic grasping, dealing with more CL learning task setting such as anthropomorphic grasping for different purposes or with different hands, and further exploration on real-world verification of actual robot platforms.

REFERENCES

[1] Davide Abati et al. "Conditional channel gated networks for task-aware continual learning". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 3931–3940.

TABLE V
COMPARISON RESULTS ON GRASPING UNSEEN OBJECTS FROM EGAD! IN SIMULATION. THE BEST RESULTS ARE REPORTED IN BOLD FONT

Methods	Penetration						Success Rate(%) ↑		
	Depth(cm) ↓			Volume(cm^3) ↓					
Finetune	0.668			5.526			45.2		
IID-Offline	0.562			6.951			73.0		
EWC	0.934			9.746			42.8		
SI	0.535			2.654			44.6		
BufferSize$\|\mathcal{M}\|$	1K	5K	10K	1K	5K	10K	1K	5K	10K
ER	1.123	0.677	**0.474**	17.455	5.654	**3.092**	51.9	65.4	74.0
ER-SI(ours)	1.019	**0.590**	0.494	14.569	**4.075**	3.265	56.0	69.4	74.5
ER-DA(ours)	0.617	0.631	0.660	5.732	6.166	6.943	78.8	77.4	75.4
ER-DA-SI(ours)	**0.597**	0.631	0.623	**5.559**	6.034	6.215	**84.0**	**78.1**	**77.6**

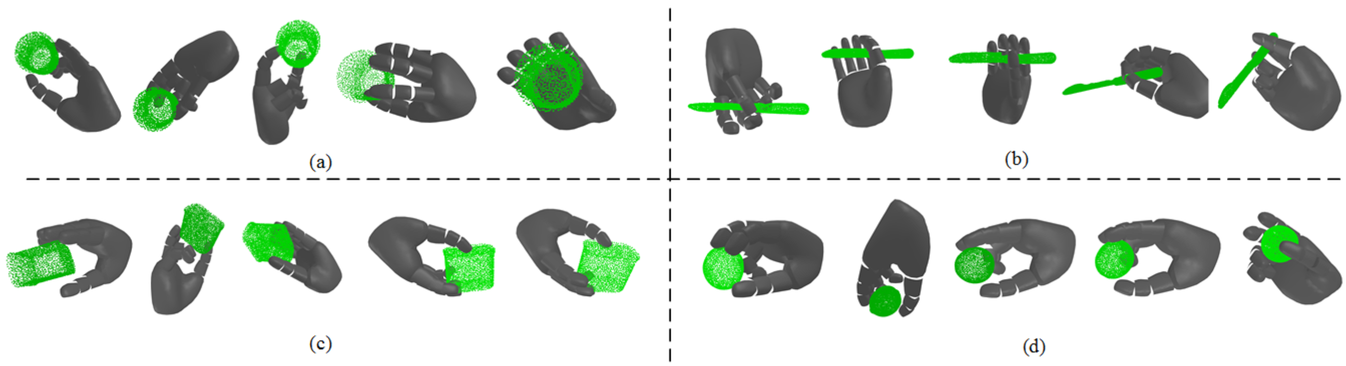


Fig. 7. Qualitative grasps generated by ER-DA-SI with 10K replay buffer on 4 objects from YCB object set. (a). Grasps for object ycb_065-f_cups_scaled. (b).Grasps for object ycb_032. (c). Grasps for object ycb_010_potted_meat_can_scaled. (d). Grasps for object ycb_014 .

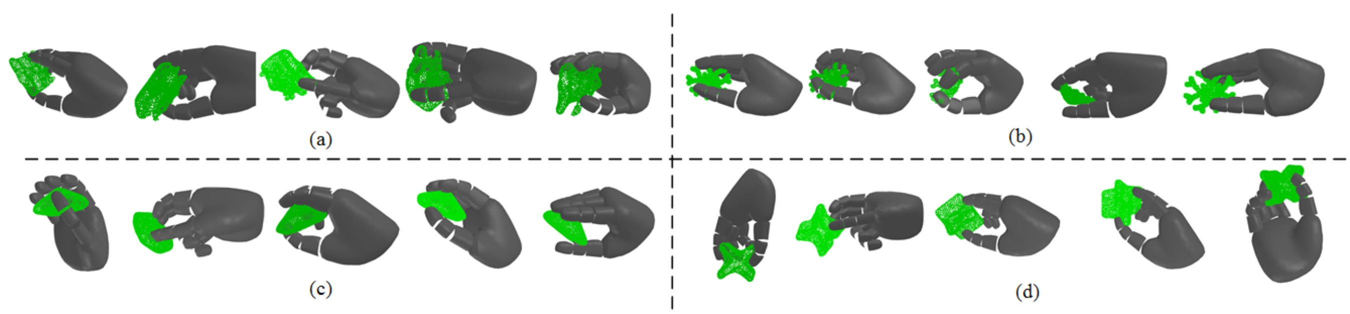


Fig. 8. Qualitative grasps generated by ER-DA-SI with 10K replay buffer on 4 objects from EGAD! object set. (a). Grasps for object G5. (b).Grasps for object G6. (c). Grasps for object G0. (d). Grasps for object A1.

[2] Umit Rusen Aktas et al. "Deep dexterous grasping of novel objects from a single view". In: *arXiv preprint arXiv:1908.04293* (2019).

[3] Rahaf Aljundi et al. "Gradient based sample selection for online continual learning". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach et al. Vol. 32. Curran Associates, Inc., 2019.

[4] Rahaf Aljundi et al. "Online continual learning with maximal interfered retrieval". In: *Advances in neural information processing systems* 32 (2019).

[5] Jihwan Bang et al. "Rainbow Memory: Continual Learning With a Memory of Diverse Samples". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2021, pp. 8218–8227.

[6] Antonio Bicchi. "On the Closure Properties of Robotic Grasping". In: *The International Journal of Robotics Research* 14.4 (1995), pp. 319–334.

[7] Jeannette Bohg et al. "Data-driven grasp synthesis—a survey". In: *IEEE Transactions on robotics* 30.2 (2013), pp. 289–309.

[8] Lucas Caccia et al. "Online learned continual compression with adaptive quantization modules". In: *International Conference on Machine Learning*. PMLR. 2020, pp. 1240–1250.

[9] Arslan Chaudhry et al. "On tiny episodic memories in continual learning". In: *arXiv preprint arXiv:1902.10486* (2019).

[10] Aristotelis Chrysakis and Marie-Francine Moens. "Online continual learning from imbalanced data". In: *In-*

ternational Conference on Machine Learning. PMLR. 2020, pp. 1952–1961.

[11] Haonan Duan et al. "Robotics Dexterous Grasping: The Methods Based on Point Cloud and Deep Learning". In: *Frontiers in Neurorobotics* 15 (2021), p. 73.

[12] Chrisantha Fernando et al. "Pathnet: Evolution channels gradient descent in super neural networks". In: *arXiv preprint arXiv:1701.08734* (2017).

[13] F Ficuciello et al. "Vision-based grasp learning of an anthropomorphic hand-arm system in a synergy-based control framework". In: *Science robotics* 4.26 (2019), eaao4900.

[14] Ian Goodfellow et al. "Generative adversarial nets". In: *Advances in neural information processing systems* 27 (2014).

[15] Saisubramaniam Gopalakrishnan et al. "Knowledge Capture and Replay for Continual Learning". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2022, pp. 10–18.

[16] Raia Hadsell et al. "Embracing change: Continual learning in deep neural networks". In: *Trends in cognitive sciences* 24.12 (2020), pp. 1028–1040.

[17] Yana Hasson et al. "Learning joint reconstruction of hands and manipulated objects". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019, pp. 11807–11816.

[18] Tyler L Hayes et al. "Remind your neural network to prevent catastrophic forgetting". In: *European Conference on Computer Vision*. Springer. 2020, pp. 466–483.

[19] Geoffrey Hinton, Oriol Vinyals, Jeff Dean, et al. "Distilling the knowledge in a neural network". In: *arXiv preprint arXiv:1503.02531* 2.7 (2015).

[20] Joel Jang et al. "Towards Continual Knowledge Learning of Language Models". In: *arXiv preprint arXiv:2110.03215* (2021).

[21] Hanwen Jiang et al. "Hand-object contact consistency reasoning for human grasps generation". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021, pp. 11107–11116.

[22] MH Jin et al. "Multisensory five-finger dexterous hand: The DLR/HIT Hand II". In: *2008 IEEE/RSJ international conference on intelligent robots and systems*. IEEE. 2008, pp. 3692–3697.

[23] James Kirkpatrick et al. "Overcoming catastrophic forgetting in neural networks". In: *Proceedings of the national academy of sciences* 114.13 (2017), pp. 3521–3526.

[24] Kilian Kleeberger et al. "A survey on learning-based robotic grasping". In: *Current Robotics Reports* 1.4 (2020), pp. 239–249.

[25] Marek S Kopicki, Dominik Belter, and Jeremy L Wyatt. "Learning better generative models for dexterous, single-view grasping of novel objects". In: *The International Journal of Robotics Research* 38.10-11 (2019), pp. 1246–1267.

[26] Qicheng Lao et al. "Continuous domain adaptation with variational domain-agnostic feature replay". In: *arXiv preprint arXiv:2003.04382* (2020).

[27] Timothée Lesort et al. "Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges". In: *Information fusion* 58 (2020), pp. 52–68.

[28] Zhizhong Li and Derek Hoiem. "Learning without forgetting". In: *IEEE transactions on pattern analysis and machine intelligence* 40.12 (2017), pp. 2935–2947.

[29] Hongzhuo Liang et al. "Multifingered Grasping Based on Multimodal Reinforcement Learning". In: *IEEE Robotics and Automation Letters* (2021).

[30] Min Liu et al. *Deep Differentiable Grasp Planner for High-DOF Grippers*. 2020.

[31] Min Liu et al. "Generating grasp poses for a high-dof gripper using neural networks". In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2019, pp. 1518–1525.

[32] Vincenzo Lomonaco et al. "Avalanche: an end-to-end library for continual learning". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021, pp. 3600–3610.

[33] Qingkai Lu et al. "Multifingered grasp planning via inference in deep neural networks: Outperforming sampling by learning differentiable models". In: *IEEE Robotics & Automation Magazine* 27.2 (2020), pp. 55–65.

[34] Arun Mallya and Svetlana Lazebnik. "Packnet: Adding multiple tasks to a single network by iterative pruning". In: *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 2018, pp. 7765–7773.

[35] Priyanka Mandikal and Kristen Grauman. "Learning dexterous grasping with object-centric visual affordances". In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2021, pp. 6169–6176.

[36] Arsalan Mousavian, Clemens Eppner, and Dieter Fox. "6-dof graspnet: Variational grasp generation for object manipulation". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019, pp. 2901–2910.

[37] German I Parisi et al. "Continual lifelong learning with neural networks: A review". In: *Neural Networks* 113 (2019), pp. 54–71.

[38] Deirdre Quillen et al. "Deep reinforcement learning for vision-based robotic grasping: A simulated comparative evaluation of off-policy methods". In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2018, pp. 6284–6291.

[39] Jathushan Rajasegaran et al. "Random path selection for incremental learning". In: *Advances in Neural Information Processing Systems* (2019).

[40] Andrei A Rusu et al. "Progressive neural networks". In: *arXiv preprint arXiv:1606.04671* (2016).

[41] Schunk. *5-finger servo-electric gripping hand SVH*. Web Page. URL: https://schunk.com/us/en/gripping-systems/special-gripper/svh/c/PGR_3161.

[42] ShadowRobot. *ShadowRobot Dexterous Hand*. Web Page. 2005. URL: https://www.shadowrobot.com/products/dexterous-hand/.

[43] Khadija Shaheen et al. "Continual Learning for Real-World Autonomous Systems: Algorithms, Challenges and Frameworks". In: *Journal of Intelligent & Robotic Systems* 105.1 (2022), p. 9. ISSN: 1573-0409. URL: https://doi.org/10.1007/s10846-022-01603-6.

[44] Hanul Shin et al. "Continual learning with deep generative replay". In: *Advances in neural information processing systems* 30 (2017).

[45] Shengyang Sun et al. "Information-theoretic Online Memory Selection for Continual Learning". In: *arXiv preprint arXiv:2204.04763* (2022).

[46] Emanuel Todorov, Tom Erez, and Yuval Tassa. "Mujoco: A physics engine for model-based control". In: *2012 IEEE/RSJ international conference on intelligent robots and systems*. IEEE. 2012, pp. 5026–5033.

[47] Gido M van de Ven, Hava T Siegelmann, and Andreas S Tolias. "Brain-inspired replay for continual learning with artificial neural networks". In: *Nature communications* 11.1 (2020), pp. 1–14.

[48] Jeffrey S Vitter. "Random sampling with a reservoir". In: *ACM Transactions on Mathematical Software (TOMS)* 11.1 (1985), pp. 37–57.

[49] Wei Wei et al. "DVGG: Deep Variational Grasp Generation for Dextrous Manipulation". In: *IEEE Robotics and Automation Letters* (2022).

[50] Bohan Wu et al. "Generative Attention Learning: a "GenerAL" framework for high-performance multi-fingered grasping in clutter". In: *Autonomous Robots* 44.6 (2020), pp. 971–990.

[51] Friedemann Zenke, Ben Poole, and Surya Ganguli. "Continual learning through synaptic intelligence". In: *International Conference on Machine Learning*. PMLR. 2017, pp. 3987–3995.

[52] Tianqiang Zhu et al. "Toward Human-Like Grasp: Dexterous Grasping via Semantic Representation of Object-Hand". In: *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2021, pp. 15721–15731. DOI: 10.1109/ICCV48922.2021.01545.