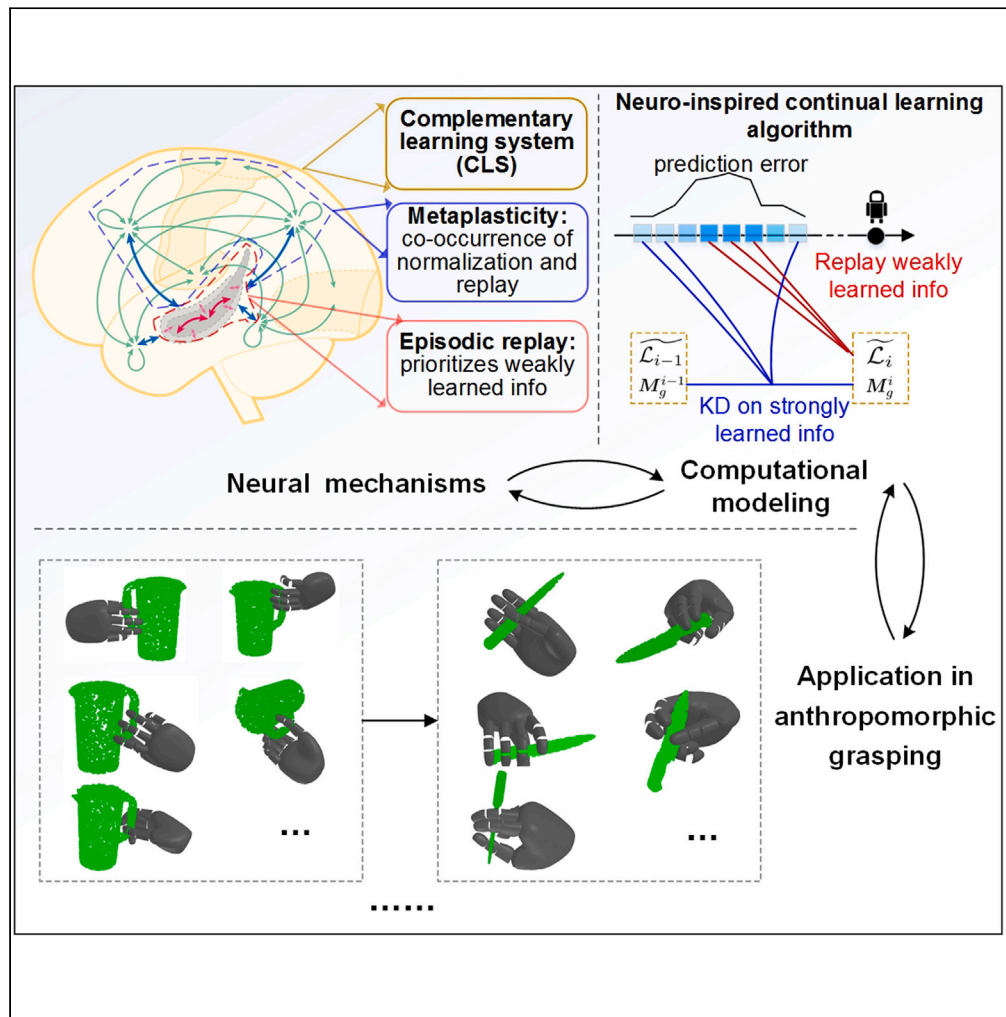


Article

Neuro-inspired continual anthropomorphic grasping



Wanyi Li, Wei Wei,
Peng Wang

peng_wang@ia.ac.cn

Highlights

Enabling anthropomorphic hand to learn grasping objects continually is critical

A neuro-inspired continual anthropomorphic grasping (NICAG) approach is proposed

NICAG performs better on alleviating forgetting and preserving grasp knowledge

NICAG has a profound impact on robots in households and factories



Article

Neuro-inspired continual anthropomorphic grasping

Wanyi Li,¹ Wei Wei,^{1,2} and Peng Wang^{1,3,4,5,*}

SUMMARY

Humans can learn continuously grasping various objects dexterously. This ability is enabled partly by underlying neural mechanisms. Most current works of anthropomorphic robotic grasping learning lack the capability of continual learning (CL). They utilize large datasets to train grasp models and the trained models are difficult to improve incrementally. By incorporating several discovered neural mechanisms supporting CL, we propose a neuro-inspired continual anthropomorphic grasping (NICAG) approach. It consists of a CL framework of anthropomorphic grasping and a neuro-inspired CL algorithm. Compared with other methods, our NICAG approach achieves better CL capability with lower loss and forgetting, and gets higher grasping success rate. It indicates that our approach performs better on alleviating forgetting and preserving grasp knowledge. The proposed system offers an approach for endowing anthropomorphic robotic hands with the ability to learn grasping objects continually and has great potential to make a profound impact on robots in households and factories.

INTRODUCTION

Humans are able to learn continuously grasping various objects dexterously throughout their lifetime under an endless variety of ever-changing scenarios. This ability is from two aspects. One is the exquisite flexibility and precision of the human hand. Another is the continual learning (CL) capability of the human brain, based on which, the dexterous grasping skill is acquired during childhood and further refined throughout life.

Biologists have tried to identify a number of underlying mechanisms that support CL. Some typical biological mechanisms include: complementary learning system (CLS),¹ episodic replay,² and meta-plasticity.³ CLS theory¹ holds that two learning systems are possessed by mammals, i.e., the hippocampus system and the neocortex system. The first allows for the rapid learning of the specifics of individual experiences which will, in turn, be played back over time to the second for acquiring structured knowledge gradually. Replay is the reactivation of neuronal activity patterns, in which neural patterns that had previously occurred during waking re-occur during later rest or sleep.² Replay appears in the hippocampus and neocortical areas, is selective and partial, and benefits subsequent memory. Schapiro et al.⁴ suggest that human hippocampal replay during rest prioritizes weakly learned information. Metaplasticity is the ability of a synapse to be modified depending on its internal biochemical states,³ which then depends on the history of synaptic modifications and recent neural activity. An instantiation of metaplasticity is regularization or normalization, with which consolidated knowledge can be protected from forgetting through synapses with a cascade of states yielding different levels of plasticity.⁵ Especially, in biological networks, normalization and synaptic changes co-occur with replay.⁶

Anthropomorphic grasping is a critical skill for robotics because robots generally need to grasp an object in the majority of manipulation tasks.^{7,8} For a robot in an open and dynamical environment, it is necessary to learn new knowledge continually over time, as it is impossible to pre-program everything in advance. The capability to learn skills and knowledge over time without forgetting the previously learned is referred to as CL.⁵ Endowing an anthropomorphic hand with the ability to learn grasping continually could have an enormous societal impact. Examples include providing assistance in the household of disabled or elder people and resorting and packaging varied goods in factories.

Existing learning-based anthropomorphic robotic grasping approaches utilize supervised learning or reinforcement learning paradigms, and train the grasping policy with large amounts of annotated data. Grasp annotations of the training data are collected by humans,⁹ with simulation¹⁰ or physical robot tests.¹¹ Given enough data, learning-based approaches achieved astonishing grasp ability. Nevertheless, they use large

¹State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

²School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

³CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China

⁴Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science and Innovation, Chinese Academy of Sciences, Hong Kong 999077, China

⁵Lead contact

*Correspondence:

peng_wang@ia.ac.cn

<https://doi.org/10.1016/j.isci.2023.106735>



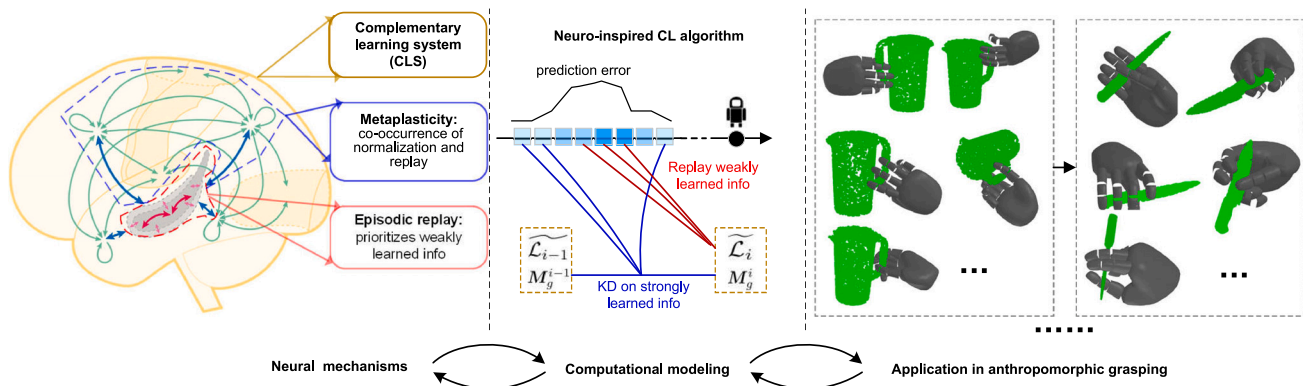


Figure 1. Schematic of the design process of neuro-inspired continual anthropomorphic grasping (NICAG)

(Left) The mechanisms concerning lifelong learning, on which we focused. The picture for lateral view of one hemisphere of the brain is adapted from Kumaran et al.¹² (Middle) Computational modeling incorporates focused mechanism into a neuro-inspired continual learning (CL) algorithm. The correspondence between specific neural mechanism and its computational component is indicated with same color, such as red is for episodic replay. (Right) Applying the designed CL algorithm into anthropomorphic grasping.

fixed-prepared dataset to train the grasp model and do not generalize well to novel objects. Furthermore, the trained grasping models are difficult to improve continually and incrementally over time.

Here, we proposed a neuro-inspired continual anthropomorphic grasping (NICAG) approach that integrates and adopts several discovered biological neural mechanisms supporting continual lifelong learning, i.e., CLS, episodic replay, and meta-plasticity. The proposed NICAG approach consists of a CL framework for anthropomorphic grasping and a neuro-inspired CL algorithm. Three layers, i.e., data layer, algorithm layer, and application layer, are included in the CL framework, thus making CL of anthropomorphic grasping possible. The neuro-inspired CL algorithm prevents forgetting and preserves grasp knowledge by replaying weakly learned information and knowledge distillation on strongly learned information, consequently, the anthropomorphic robotic hands can learn to grasp different objects continually and incrementally over long sequential grasp stream. We validate the proposed approach through dataset experiments and simulated experiments. Compared with other methods, our NICAG approach not only achieves better CL capability with lower average loss and forgetting but also gets a higher success rate (SR) for grasping. It indicates that our approach performs better on alleviating forgetting and preserving grasp knowledge. The proposed system offers an approach for endowing anthropomorphic robotic hands with the ability to learn to grasp different objects continually and incrementally over time and has great potential to make a profound impact on robots in households and factories. The contributions of this paper are as follows.

1. A NICAG approach incorporates several discovered biological mechanisms of lifelong learning.
2. A CL framework of anthropomorphic grasping that includes data layer, algorithm layer, and application layer. It makes CL of anthropomorphic grasping possible.
3. A neuro-inspired CL algorithm that prevents forgetting and preserves grasp knowledge by replaying weakly learned information and knowledge distillation on strongly learned information, thus the anthropomorphic robotic hands can learn to grasp different objects continually and incrementally over long sequential grasp stream.
4. A validation of the proposed NICAG approach through dataset experiments and simulated experiments, demonstrating our method achieves better performance than state-of-the-art CL methods. It not only achieves better CL capability with lower average loss and forgetting but also gets higher SR for grasping.

RESULTS

The three focused biological neural mechanisms supporting continual lifelong learning, i.e., CLS, episodic replay, and meta-plasticity, have been well described in ref.^{1,2,3,4,6} We propose our NICAG approach based on these mechanisms. The design process of our NICAG approach is shown in Figure 1. We will describe the CL framework of anthropomorphic grasping, the neuro-inspired CL algorithm and experimental results in the following.

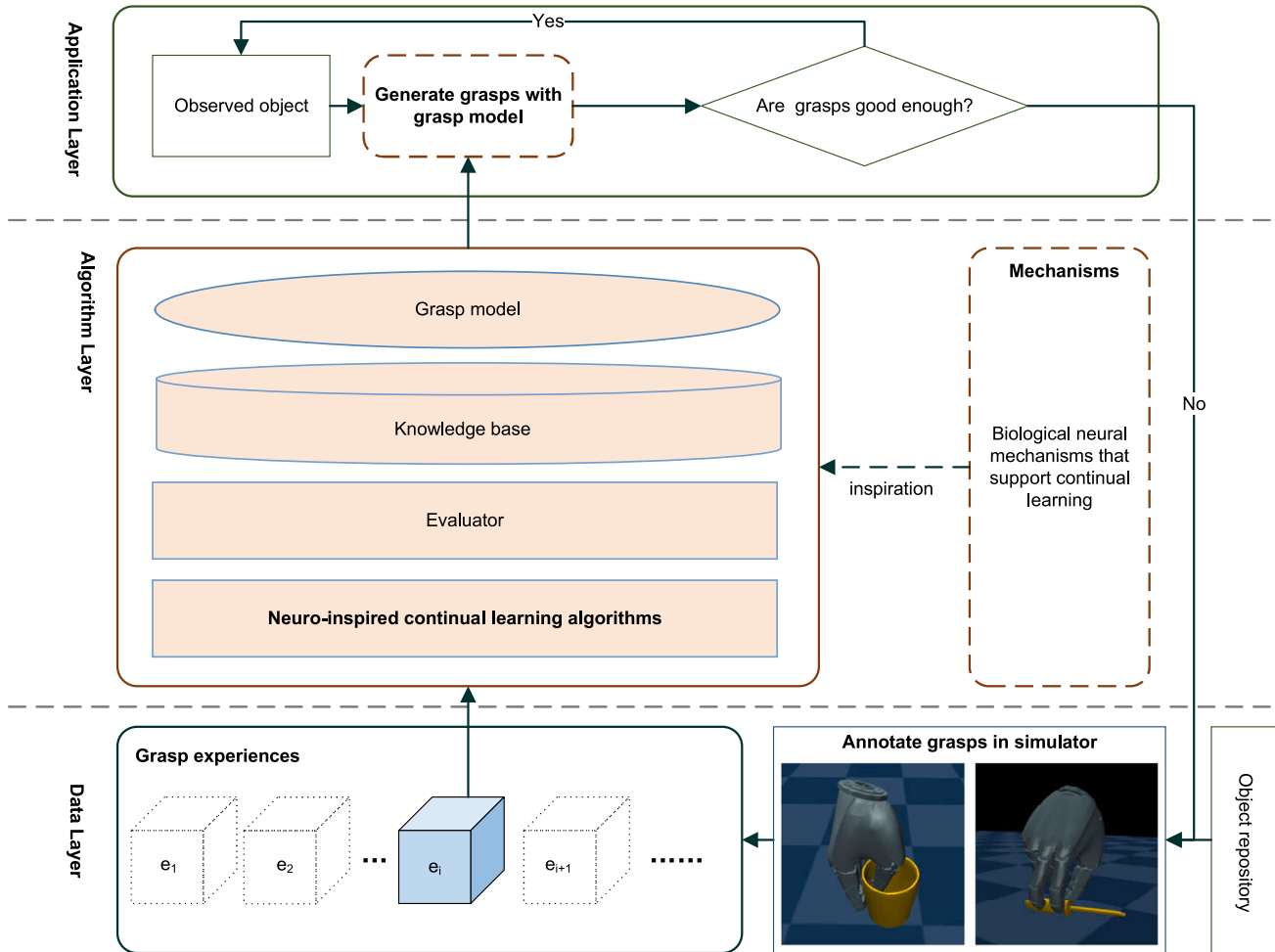


Figure 2. Continual learning framework of anthropomorphic grasping

The problem and the continual learning framework of anthropomorphic grasping

To enable anthropomorphic robotic hands to learn to grasp objects continually over time, we design a neuro-inspired CL framework for anthropomorphic grasping (NICAG-framework), which is shown in Figure 2. There are three layers in the NICAG-framework: data layer, algorithm layer, and application layer. Data layer is responsible for generating the stream of anthropomorphic grasping experiences. Algorithm layer trains the grasp model based on information from data layer. Application layer applies trained grasp model to objects in the field, those objects with bad grasps are sent to the data layer for better learning.

Formally, in the NICAG-framework, a CL algorithm A_{CL} is expected to update its internal state, e.g., its internal grasp model M_g and knowledge base representing as specific data structures, based on a non-stationary sequentially accessible stream of anthropomorphic grasping experiences $\mathbf{E} = (e_1, e_2, \dots, e_i, \dots, e_n)$. The objective of A_{CL} is to improve its performance on a set of grasp metrics (p_1, \dots, p_m) as evaluated on a test stream of experiences (e_1^t, \dots, e_n^t) .

With respect to the stream of anthropomorphic grasping experiences $\mathbf{E} = (e_1, e_2, \dots, e_i, \dots, e_n)$, the i -th experience consists of $e_i = \{(\mathcal{P}_k, \mathbf{g}_k)_{k=1}^{n_i}\}$, where each pair constitutes a grasp example of a point cloud \mathcal{P}_k of the observed object, and a grasp \mathbf{g}_k . The grasp is defined as $\mathbf{g} = \{\mathbf{p}, \boldsymbol{\theta}\}$. The hand wrist pose \mathbf{p} is given in special Euclidean group $SE(3)$, 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. In this work we use anthropomorphic robot hand DLR/HIT Hand II, of which $\boldsymbol{\theta} \in \mathbb{R}^{20}$.

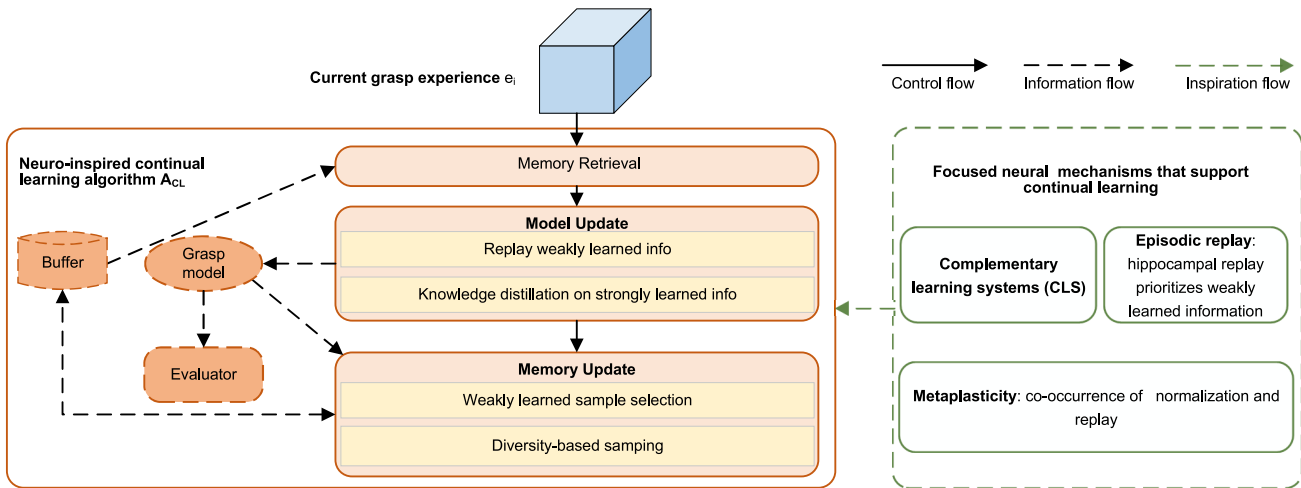


Figure 3. Schematic view of the proposed method

Within the learning framework, this paper proposes a neuro-inspired CL algorithm A_{CL} . It is used to update internal grasp model M_g and knowledge base. The grasp model M_g takes point cloud \mathcal{P} of the observed object as input and predicts high quality grasps. The knowledge base is represented as a memory buffer \mathcal{M} and a last trained model. The details of the neuro-inspired CL algorithm A_{CL} are provided in the next subsection.

The neuro-inspired continual learning algorithm

We first describe our neuro-inspired CL algorithm A_{CL} at a high level here. A_{CL} prevents forgetting and preserves grasp knowledge in both sample-space and function-space. Replaying weakly learned information is for preserving grasp knowledge in sample space, while knowledge distillation on strongly learned information is for keeping knowledge in function-space. The schematic view of the proposed algorithm is given in Figure 3. It consists of three major steps: memory retrieval based on learnability criterion, model update by replay weakly learned information and knowledge distillation on strongly learned information, and memory update with weakly learned sample selection and diversity-based sampling. In the following subsections, we first introduce the learnability criterion, and then provide details of three major steps.

Learnability criterion

To indicate whether a training sample is strongly learned or weakly learned, we adopt the learnability criterion from Sun et al.,¹³ which measures how much the grasp model M_g can explain the training sample $(\mathcal{P}_\star, \mathbf{g}_\star)$ once it has absorbed its information in the memory. Adapted from Sun et al.,¹³ we define the learnability as follows:

$$s_{learn}(\langle \mathcal{P}_\star, \mathbf{g}_\star \rangle; \mathcal{M}) = \log p(\mathbf{g}'_\star | \mathbf{g}_\star, \mathcal{M}; \mathcal{P}_\star, \mathcal{P}_M) |_{\mathbf{g}'_\star = \mathbf{g}_\star} \quad (\text{Equation 1})$$

where, \mathbf{g}'_\star and \mathbf{g}_\star are used to represent two realizations of the same random variable describing a grasp.

Due to the used grasp model is a variational autoencoder (VAE) based generative model, after the grasp model M_g have visited a grasp experience, the learnability of sample $(\mathcal{P}_\star, \mathbf{g}_\star)$ with respect to M_g is calculated specifically as a quantity related with its loss:

$$s_{learn}(\langle \mathcal{P}_\star, \mathbf{g}_\star \rangle, M_g) = e^{-\mathcal{L}(\langle \mathcal{P}_\star, \mathbf{g}_\star \rangle, M_g)} \quad (\text{Equation 2})$$

where $\mathcal{L}(\langle \mathcal{P}_\star, \mathbf{g}_\star \rangle, M_g)$ is the loss of sample $(\mathcal{P}_\star, \mathbf{g}_\star)$ after passing through the grasp model M_g . The information of memory buffer \mathcal{M} in Equation 1 has been incorporated into the grasp model M_g during the training process on the latest visited grasp experience.

Memory retrieval based on learnability criterion

Based on learnability score, there are two parts in memory buffer \mathcal{M} : weakly learned samples \mathcal{S}_{wl} and diverse strongly learned samples \mathcal{S}_{sj} . They are with same size $\frac{|\mathcal{M}|}{2}$, where $|\mathcal{M}|$ is the size of memory buffer \mathcal{M} . The memory buffer is updated once an experience was learned, details are in Section "memory update with weakly learned sample selection and diversity-based sampling". For each incoming mini-batch B_j

drawn from current grasp experience e_i , the memory retrieval step randomly selects $\frac{|B_j|}{2}$ weakly learned samples (denoted as B_j^{wl}) and $\frac{|B_j|}{2}$ diverse strongly learned samples (denoted as B_j^{sl}) from the memory buffer \mathcal{M} , $|B_j|$ is the batch size. To enhance the diversity of retrieved samples, 3D data augmentation on B_j^{wl} and B_j^{sl} are applied. Operations for 3D data augmentation include jitter, dropout and rotation. Jitter operation adds a clipped Gaussian noise with zero mean and standard deviation σ to the position of each point. Dropout augmentation throw away points randomly with max ratio r_{max} . And rotation augmentation randomly rotates the object \mathcal{P}_k and grasp \mathbf{g}_k along three axes. The retrieved and augmented samples are used for updating grasp model M_g , as described in Section "model update by replay weakly learned information and knowledge distillation on strongly learned information".

Model update by replay weakly learned information and knowledge distillation on strongly learned information

To preserve grasp knowledge in sample space, we replay the retrieved weakly learned samples B_j^{wl} , the loss of weakly learned information replay is defined as:

$$\mathcal{L}_{replay} = \mathcal{L}(B_j^{wl}, M_g) \quad (\text{Equation 3})$$

To keep grasp knowledge in function-space, we apply knowledge distillation on strongly learned samples B_j^{sl} . With respect to the strongly learned samples, it is expected that the current grasp model M_g^i and the previous grasp model M_g^{i-1} encodes latent code and generate final grasp in the same way. We utilize a KL-divergence loss to enforce the latent code distributions of M_g^i and M_g^{i-1} to be close, and use a reconstruction loss to encourage the output of M_g^i and M_g^{i-1} to be same. The KL-divergence loss and the reconstruction loss are formulated in Equations 4 and 5. The loss of knowledge distillation consist of two terms, i.e., the KL-divergence loss and the reconstruction loss, and is defined in Equation 6.

$$\mathcal{L}_{kl} = KL(Q(E_i(B_j^{sl})) || P(E_{i-1}(B_j^{sl}))) \quad (\text{Equation 4})$$

where E_i and E_{i-1} are the encoders of current grasp model M_g^i and the previous grasp model M_g^{i-1} , respectively, $KL(Q || P)$ is the Kullback-Leibler (KL) divergence to measure how different these two distributions Q and P are.

$$\mathcal{L}_r = \lambda_v \frac{1}{N} \sum_{k=1}^N \left\| M_g^i(B_j^{sl})_k^V - M_g^{i-1}(B_j^{sl})_k^V \right\|_2^2 + \lambda_\theta \cdot \left\| M_g^i(B_j^{sl})^\theta - M_g^{i-1}(B_j^{sl})^\theta \right\|_1 \quad (\text{Equation 5})$$

The reconstruction loss is based on the reconstructed hand mesh. It consists of two terms: hand mesh vertices displacement and joint angles error. In Equation 5, $M_g^i(B_j^{sl})^V$ is the vertices set of the reconstructed hand mesh, and $M_g^i(B_j^{sl})^\theta$ is the joint angle of generated grasps.

$$\mathcal{L}_{kd} = \mathcal{L}_{kl}(B_j^{sl}, M_g^{i-1}, M_g^i) = \lambda_{kl} \cdot \mathcal{L}_{kl} + \mathcal{L}_r \quad (\text{Equation 6})$$

where M_g^{i-1} is the previous version of grasp model, parameters of which just were updated based on the last experience e_{i-1} , M_g^i is the current version of grasp model which is visiting the current experience e_i . λ_{kl} of Equation 6, λ_v and λ_θ of Equation 5 are constants to balance the losses.

Combining with the loss of mini-batch B_j drawn from current grasp experience e_i , loss of weakly learned replay, and loss of knowledge distillation on strongly learned information, we perform model update by optimizing the following loss with respect to the parameters of the grasp model M_g :

$$\mathcal{L}_{grasp}^{cl} = \mathcal{L}(B_j, M_g) + \lambda_{replay} \cdot \mathcal{L}_{replay} + \mathcal{L}_{kd} \quad (\text{Equation 7})$$

Memory update with weakly learned sample selection and diversity-based sampling

After the grasp model M_g has been trained on a grasp experience, we perform the memory update step. Firstly, we merge the samples from memory buffer \mathcal{M} and samples from the current experience e_i as $\mathcal{S}_{\mathcal{M}} \cup \mathcal{S}_{e_i}$ and calculate their learnability scores according to Equation 1. Secondly, from $\mathcal{S}_{\mathcal{M}} \cup \mathcal{S}_{e_i}$, we select top $\frac{|\mathcal{M}|}{2}$ samples with lowest learnability score as weakly learned samples, noted as \mathcal{S}_{wl} . And then, we select $\frac{|\mathcal{M}|}{2}$ examples from rested samples $\mathcal{S}_{rest} = \mathcal{S}_{\mathcal{M}} \cup \mathcal{S}_{e_i} \setminus \mathcal{S}_{wl}$ descending by learnability score with an interval of $|\mathcal{S}_{rest}| / \left(\frac{|\mathcal{M}|}{2}\right)$, as a result, $\frac{|\mathcal{M}|}{2}$ diverse strongly learned samples are included, which is denoted as \mathcal{S}_{sl} . Finally, the contents of memory buffer is replaced with the selected $|\mathcal{M}|$ samples, i.e., \mathcal{S}_{wl} and \mathcal{S}_{sl} .

Table 1. Hyperparameters for the compared methods

Methods	Hyperparameter grid	Tuned hyperparameter
Fine-tune	— ^a	—
IID-Offline	—	—
EWC	$\lambda : [0.1, 1, 100]$	$\lambda = 0.1$
SI	$c : [0.1, 0.5, 1]$	$c = 0.5$
ER	$\lambda_{replay} : [0.1, 0.5, 1]$	$\lambda_{replay} = 0.5$
ER-RM	$\lambda_{replay} : [0.1, 0.5, 1]$	$\lambda_{replay} = 0.5$
NI-WL ^b	$\lambda_{replay} : [0.1, 0.5, 1]$	$\lambda_{replay} = 0.5$
NI-WL-RD ^b	$\lambda_{replay} : [0.1, 0.5, 1]$	$\lambda_{replay} = 0.5$
NI-WL-RM ^b	$\lambda_{replay} : [0.1, 0.5, 1]$	$\lambda_{replay} = 0.5$
NI-WL-RM-KD ^b	$\lambda_{replay} : [0.1, 0.5, 1], \lambda_{kl}, \lambda_{\theta} : [0.1, 0.5, 1], \lambda_{\nu} : [1, 10, 30]$	$\lambda_{replay} = 0.5, \lambda_{kl} = 0.1, \lambda_{\theta} = 0.5, \lambda_{\nu} = 30$

^a“—” means not applicable.

^bNI-WL, NI-WL-RD, NI-WL-RM, and NI-WL-RM-KD are four variants of our proposed method.

Experimental results

A validation of the proposed NICAG approach through dataset experiments and simulated experiments, demonstrating our method achieves better performance than state-of-the-art CL methods. It not only achieves better CL capability with lower average loss and forgetting but also gets higher SR for grasping.

To validate our proposed NICAG approach, we conduct experiments on both dataset and in simulation against a set of CL methods. The experiments aim to evaluate the CL capability and the grasp performance of the NICAG approach.

For evaluation metrics in dataset experiments, Average Loss (L_i) and Average Forgetting (F) are used to evaluate the CL capability of compared methods. The Average Loss (L_i) is the averaged loss of grasp model on test sets of learned experiences (e_1, e_2, \dots, 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, \dots, e_n) that is defined in Equation 8. The Average Forgetting (F) is defined upon Average Loss in Equation 9.

$$mAL = \frac{1}{n} \sum_{i=1}^n L_i \quad (\text{Equation 8})$$

$$F = \frac{1}{n-1} \sum_{i=2}^n \min(L_i - L_{i-1}, 0) \quad (\text{Equation 9})$$

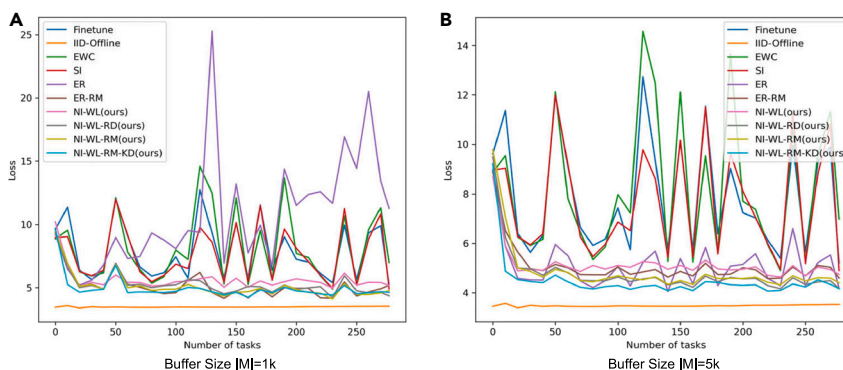


Figure 4. The average loss on test set without random rotation so far measured by the end of each task (object)

(A) 1K buffer size is used for replay related methods.

(B) 5K buffer size is used for replay related methods.

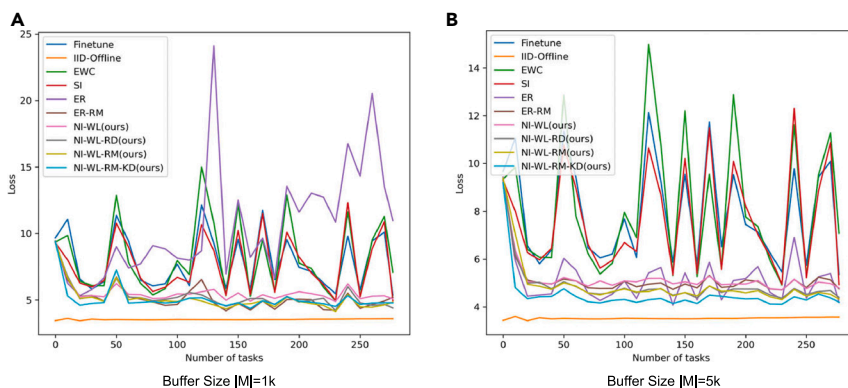


Figure 5. The average loss on test set with random rotation so far measured by the end of each task (object)
(A) 1K buffer size is used for replay related methods.
(B) 5K buffer size is used for replay related methods.

We use three quantitative metrics for evaluation in simulation: Success Rate (SR), Penetration Depth and Penetration Volume between the hand mesh and the target object. The three used metrics keep consistent with previous literature Mousavian et al.¹⁴ and Hasson et al.¹⁵ SR is commonly used in grasping tasks to measure the stability and quality of the generated grasps. For penetration depth and penetration volume, the implementation of Jiang et al.¹⁶ is used. 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.

We describe the implementation details of all compared CL algorithms here. All CL algorithms are implemented using Avalanche,¹⁷ which is an end-to-end CL library based on PyTorch. For Ild-Offline, i.e., the variational grasp generator in DVGG¹⁰ (deep variational grasp generation), we use the implementation of Wei et al.¹⁰ For training of the 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 an RTX-3090 GPU. We present the detailed hyperparameters in Table 1.

Results on dataset

To evaluate the CL capability of the proposed neuro-inspired algorithm, we compare it with other six methods on dataset. Details of the dataset are described in Section "dataset" of STAR Methods. The compared methods include four typical CL methods, namely, elastic weight consolidation (EWC),¹⁸ synaptic intelligence (SI),¹⁹ experience replay (ER),²⁰ rainbow memory (ER-RM),²¹ and two baselines, i.e., Fine-tune and Ild-Offline. Our proposed neuro-inspired CL algorithm includes four variants: NI-WL is the weakly learned replay, NI-WL-RD is the integration of weakly learned replay and randomly selective ER, NI-WL-RM is the integration of weakly learned replay and rainbow memory replay, NI-WL-RM-KD is the integration of weakly learned replay and knowledge distillation on strongly learned information. The description of compared methods is in Section "description of compared methods" of STAR Methods. We will report and analyze the evolution of test loss and forgetting along with training, the mAL and average forgetting, and the loss on the combined test set when the grasp model is finally trained on all grasp experiences, respectively.

Evolution of test loss and forgetting along with training. In Figures 4 and 5, we show how the loss on test set so far evolve along with seeing more tasks, i.e., seeing more objects. Figure 4 is for the version of which test set is without random rotation, while Figure 5 is for that with random rotation. The evolution processes of test forgetting along with training are shown in Figure 6 (without random rotation) and Figure 7 (with random rotation). The lower and smoother the loss is, the better the corresponding CL method is. So is the forgetting. From left to right in Figures 4, 5, 6, and 7, memory size changes with 1K–5K. As shown in Figures 4, 5, 6, and 7, the naive fine-tune has high loss (also high forgetting) and oscillates up and down with a large attitude, indicating catastrophic forgetting occurs. EWC is even worse than Finetune under 5K buffer size, due to the saturation-prone property of regularization methods in the long steam. SI is better than Finetune, but is still with high loss, high forgetting, and large oscillation. ER has high loss, high forgetting, and is 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, low forgetting, and small oscillation when big buffer size is

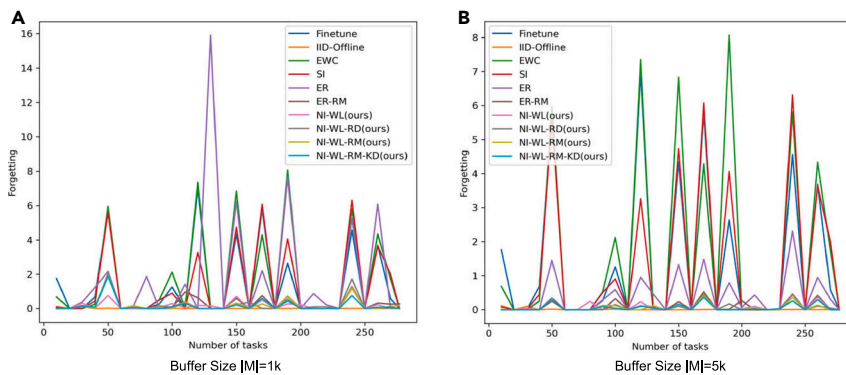


Figure 6. The forgetting metric on test set without random rotation so far measured by the end of each task (object)

(A) 1K buffer size is used for replay related methods.
(B) 5K buffer size is used for replay related methods.

used, such as with 5K memory buffer. Thanks to the diversity of the replayed samples, ER-RM performs well under different buffer size. By contrast, the variants of our proposed method perform better with lower loss, lower forgetting, and smaller oscillation. NI-WL-RM-KD achieves best results, which is very close to the IID-Offline, even with only 1K memory buffer. The visualized tendencies of Figures 4, 5, 6, and 7 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. The quantitative results of Figures 4, 5, 6, and 7 are summarized in Table 2, which is provided as the mAL and average forgetting F. Firstly, on test set without random rotation, compared to EWC, SI, ER, and ER-RM, our proposed NI-WL-RM-KD with small buffer size 1K shows $\sim 39.6\%$, $\sim 35.4\%$, $\sim 53.9\%$, and $\sim 3.5\%$ relative reduction in mAL, $\sim 90.1\%$, $\sim 87.4\%$, $\sim 90.8\%$, and $\sim 40.7\%$ reduction in forgetting, respectively. This indicates that the proposed NI-WL-RM-KD is able to work under small memory cost. Under other buffer size conditions, four variants of our proposed method, including NI-WL, NI-WL-RD, NI-WL-RM, and NI-WL-RM-KD, consistently outperform other CL methods with a large margin in terms of mAL and forgetting both. Second, a similar trend is shown both on test set without rotation and with rotation, which indicates the robustness of the proposed methods to random rotation. Further, compared to IID-Offline, all CL methods are with more 16% in mean average losses indicating gap still exists.

Loss on combined test set of finally trained model. In Table 3, we provide the average loss of finally trained models for all compared methods. The losses are calculated on combined test set or combined

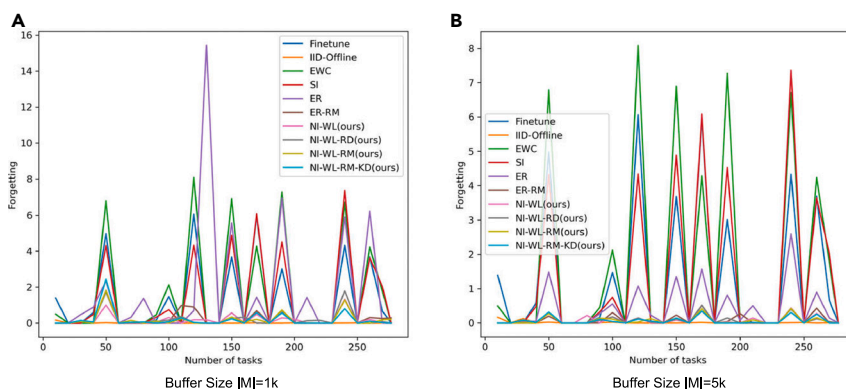


Figure 7. The forgetting metric on test set with random rotation so far measured by the end of each task (object)

(A) 1K buffer size is used for replay related methods.
(B) 5K buffer size is used for replay related methods.

Table 2. Mean average loss mAL and average forgetting F of compared methods

Methods	$mAL\downarrow$ (TeWoR) ^a	$F\downarrow$ (TeWoR)	$mAL\downarrow$ (TeWR) ^a	$F\downarrow$ (TeWR)
Fine-tune	7.889	1.368	7.924	1.284
IID-Offline	3.470	— ^b	3.520	—
EWC	8.199	1.712	8.209	1.754
SI	7.657	1.344	7.686	1.382
BufferSize $\mathcal{M} = 1K$				
ER	10.723	1.843	10.554	1.749
ER-RM	<u>5.132</u>	0.285	<u>5.169</u>	0.284
NI-WL ^c	5.630	<u>0.175</u>	5.543	<u>0.175</u>
NI-WL-RD ^c	5.255	0.215	5.243	0.208
NI-WL-RM ^c	<u>5.100</u>	<u>0.198</u>	<u>5.089</u>	<u>0.192</u>
NI-WL-RM-KD ^c	<u>4.950</u>	<u>0.169</u>	<u>5.066</u>	<u>0.194</u>
BufferSize $\mathcal{M} = 5K$				
ER	5.094	0.406	5.147	0.416
ER-RM	5.077	0.096	5.143	0.092
NI-WL	5.144	0.091	5.197	0.076
NI-WL-RD	<u>4.768</u>	<u>0.071</u>	<u>4.891</u>	<u>0.073</u>
NI-WL-RM	<u>4.855</u>	<u>0.064</u>	<u>4.887</u>	<u>0.068</u>
NI-WL-RM-KD	<u>4.490</u>	<u>0.062</u>	<u>4.511</u>	<u>0.063</u>
BufferSize $\mathcal{M} = 10K$				
ER	4.975	0.420	4.980	0.412
ER-RM	4.948	0.082	5.035	0.091
NI-WL	4.996	0.082	4.988	<u>0.079</u>
NI-WL-RD	<u>4.810</u>	<u>0.058</u>	<u>4.853</u>	<u>0.066</u>
NI-WL-RM	<u>4.558</u>	<u>0.068</u>	<u>4.574</u>	<u>0.066</u>
NI-WL-RM-KD	<u>4.029</u>	<u>0.069</u>	<u>4.045</u>	<u>0.071</u>

Bold underline, italic underline, and underline font highlights the first place, second place, and third place with same BufferSize, respectively.

^aTeWoR is short for test set without rotation, TeWR is short for test set with rotation.

^b“—” means not applicable.

^cNI-WL, NI-WL-RD, NI-WL-RM, and NI-WL-RM-KD are four variants of our proposed method.

training set of all experiences. There are four losses for each method, 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 demonstrated in Table 3, four losses of Finetune are quite high and are all above 5.25. As expected lower bound, losses of IID-Offline are low and below 3.6. Consistent with evolution of test loss along with training in Figures 4 and 5, EWC has high losses around 6.9 which are all larger than those of Finetune, losses of SI are around 4.9 and are slightly lower than Finetune's. Losses of ER appear similar tendency with Figures 4 and 5, is high under small buffer size, while is low under big buffer size. ER-RM performs better than ER due to its diversity of replay samples. For the variants of our proposed method, ER-WL-RD gets the most significant drop of loss under 1K buffer size, from 11.267 of ER to 4.356. This indicates the weakly learned plus diverse sampling enhanced the representative and diversity of replay samples. The full-armed version of our proposed method, i.e., ER-WL-RM-KD, gets the most top places.

Results in simulation

To illustrate the effectiveness of the proposed approach on continually generating anthropomorphic grasps with high quality, we conduct simulated experiments in the physics-based simulator MuJoCo.²² 58 objects from YCB (yale-cmu-berkeley) dataset²³ (seen) and 48 objects from EGAD!²⁴ (unseen) are

Table 3. Loss L on combined set (test set or training set) of finally trained models for compared methods

Methods	$L \downarrow$ (TeWoR) ^a	$L \downarrow$ (TeWR) ^a	$L \downarrow$ (TrWoR) ^a	$L \downarrow$ (TrWR) ^a
Fine-tune	5.299	5.257	5.298	5.256
IID-Offline	3.519	3.571	3.521	3.572
EWC	6.986	7.071	6.984	7.074
SI	4.899	4.938	4.897	4.934
BufferSize $\mathcal{M} = 1K$				
ER	11.267	10.982	11.283	10.991
ER-RM	5.186	5.195	5.187	5.192
NI-WL ^b	5.172	5.061	5.172	5.060
NI-WL-RD ^b	<u>4.356</u>	<u>4.391</u>	<u>4.358</u>	<u>4.391</u>
NI-WL-RM ^b	<u>4.766</u>	<u>4.830</u>	<u>4.769</u>	<u>4.831</u>
NI-WL-RM-KD ^b	<u>4.631</u>	<u>4.782</u>	<u>4.632</u>	<u>4.776</u>
BufferSize $\mathcal{M} = 5K$				
ER	<u>4.182</u>	<u>4.188</u>	<u>4.182</u>	<u>4.188</u>
ER-RM	4.596	4.759	4.599	4.760
NI-WL	4.740	4.790	4.737	4.792
NI-WL-RD	<u>4.172</u>	4.386	<u>4.170</u>	4.388
NI-WL-RM	4.400	<u>4.319</u>	4.402	<u>4.320</u>
NI-WL-RM-KD	<u>4.162</u>	<u>4.200</u>	<u>4.168</u>	<u>4.197</u>
BufferSize $\mathcal{M} = 10K$				
ER	<u>3.754</u>	<u>3.799</u>	<u>3.755</u>	<u>3.800</u>
ER-RM	4.616	4.707	4.614	4.704
NI-WL	4.607	4.513	4.609	4.514
NI-WL-RD	4.542	4.510	4.540	4.508
NI-WL-RM	<u>4.182</u>	<u>4.182</u>	<u>4.180</u>	<u>4.180</u>
NI-WL-RM-KD	<u>4.028</u>	<u>4.045</u>	<u>4.029</u>	<u>4.047</u>

Bold underline, italic underline, and underline font highlights the first place, second place, and third place with same BufferSize, respectively.

^aTeWoR is short for test set without rotation, TeWR is short for test set with rotation, while TrWoR is short for training set without rotation, and TrWR is short for training set with rotation.

^bNI-WL, NI-WL-RD, NI-WL-RM, and NI-WL-RM-KD are four variants of our proposed method.

used. For each object, the completed 3D point cloud is taken as the input of the trained grasp model M_g , and M_g generates 20 grasps randomly. In the simulator, we perform grasp with generated grasp configuration for all used objects and calculate the metrics, i.e., SR, Penetration Depth, and Penetration Volume between the hand mesh and the target object. The steps in the physical simulation process are described in Section "steps for the simulated experiments" of STAR Methods.

In Table 4 and Table 5, the compared results on grasping seen objects from YCB and unseen objects from EGAD! in the simulation are provided respectively. As shown in Table 4, most top places with respect to SR on grasping objects from YCB are achieved by our proposed approach. At the same time, the variants of our proposed method, NI-WL, NI-WL-RD, NI-WL-RM, and NI-WL-RM-KD, are consistently with lower Penetration. Table 5 shows a similar tendency also. Overall, the proposed approach outperforms other alternatives on grasping object in simulation with higher SR and lower penetration including depth and volume. Moreover, it is observed that ER-WL for different buffer size, ER-WL-RD, ER-WL-RM, and ER-WL-RM-KD under 5K and 10K buffer size, outperform IID-Offline for unseen EGAD! object dataset, perhaps due to the bias from the dominant objects in IID-Offline. Qualitative results shown in Figures 8 and 9 demonstrate that our proposed method is able to generate diverse reasonable grasps.

Table 4. Compared results on grasping seen objects from YCB in simulation

Methods	P-Depth ^a (cm) ↓	P-Volume ^a (cm ³) ↓	Success Rate (%) ↑
Fine-tune	0.824	9.142	32.8
IID-Offline	0.642	7.273	62.0
EWC	0.714	8.530	23.3
SI	0.733	8.663	33.6
BufferSize \mathcal{M} = 1K			
ER	1.262	23.204	41.4
ER-RM	0.895	11.928	<u>49.3</u>
NI-WL ^b	<u>0.516</u>	<u>6.115</u>	<u>54.7</u>
NI-WL-RD ^b	<u>0.755</u>	10.073	45.3
NI-WL-RM ^b	<u>0.731</u>	<u>9.963</u>	<u>50.6</u>
NI-WL-RM-KD ^b	0.818	<u>6.413</u>	47.1
BufferSize \mathcal{M} = 5K			
ER	0.668	8.021	53.0
ER-RM	<u>0.477</u>	<u>5.211</u>	<u>55.2</u>
NI-WL	<u>0.496</u>	<u>5.868</u>	<u>55.5</u>
NI-WL-RD	0.533	6.587	54.7
NI-WL-RM	<u>0.531</u>	5.908	54.5
NI-WL-RM-KD	0.613	<u>4.400</u>	<u>55.4</u>
BufferSize \mathcal{M} = 10K			
ER	0.608	6.340	<u>55.4</u>
ER-RM	<u>0.455</u>	<u>4.703</u>	<u>55.9</u>
NI-WL	0.512	5.720	52.7
NI-WL-RD	<u>0.438</u>	<u>4.458</u>	<u>54.6</u>
NI-WL-RM	<u>0.482</u>	5.201	53.1
NI-WL-RM-KD	0.661	<u>4.355</u>	51.3

Bold underline, italic underline, and underline font highlights the first place, second place, and third place with same BufferSize, respectively.

^aP-Depth is short for Penetration Depth, P-Volume is short for Penetration Volume.

^bNI-WL, NI-WL-RD, NI-WL-RM, and NI-WL-RM-KD are four variants of our proposed method.

DISCUSSION

In this work, the problem of continual anthropomorphic grasping is considered. In particular, a NICAG approach is developed, which incorporates several discovered biological neural mechanisms supporting continual lifelong learning and consists of a CL framework of anthropomorphic grasping and a neuro-inspired CL algorithm. The experiments carried out on dataset and in simulation provide encouraging results, showing that this approach achieves better CL capability with lower average loss and forgetting, but also gets higher SR for grasping, with reference to some CL metrics and grasp quality metrics. The proposed system offers an approach for endowing anthropomorphic robotic hands with the ability to learn to grasp different objects continually and incrementally over time, and has great potential to make a profound impact on robots in households and factories.

Starting from this work, some future directions are worthwhile considering. Firstly, dealing with more task settings of continual anthropomorphic grasping, one example is CL of anthropomorphic grasping for different purposes (e.g., tool use, hand over, pick, and place), or with different hands. Secondly, due to reinforcement learning could be utilized to reduce repeated failures, integrating supervised learning and reinforcement learning into the continual anthropomorphic grasping framework is also a good direction. Thirdly, developing composite continual anthropomorphic grasping systems that incorporate more biological mechanisms of lifelong learning and human grasping is of great significance. Some biological mechanisms to be incorporated include neuromodulation,²⁵ context-dependent perception and gating,²⁶

Table 5. Compared results on grasping unseen objects from EGAD! in simulation

Methods	P-Depth ^a (cm) ↓	P-Volume ^a (cm ³) ↓	Success Rate (%) ↑
Fine-tune	0.668	5.526	45.2
IID-Offline	0.562	6.951	73.0
EWC	0.446	3.341	32.2
SI	0.535	2.654	44.6
BufferSize \mathcal{M} = 1K			
ER	1.123	17.455	51.9
ER-RM	0.964	20.867	55.5
NI-WL ^b	<u>0.458</u>	<u>6.874</u>	<u>81.7</u>
NI-WL-RD ^b	<u>0.760</u>	<u>13.237</u>	65.3
NI-WL-RM ^b	0.805	16.537	<u>69.4</u>
NI-WL-RM-KD ^b	<u>0.763</u>	<u>5.822</u>	<u>69.0</u>
BufferSize \mathcal{M} = 5K			
ER	0.677	<u>5.654</u>	65.4
ER-RM	<u>0.448</u>	<u>6.593</u>	<u>78.5</u>
NI-WL	<u>0.486</u>	7.390	<u>76.9</u>
NI-WL-RD	<u>0.513</u>	8.590	<u>77.9</u>
NI-WL-RM	0.645	9.317	73.4
NI-WL-RM-KD	0.536	<u>3.686</u>	<u>76.9</u>
BufferSize \mathcal{M} = 10K			
ER	0.474	<u>3.092</u>	74.0
ER-RM	<u>0.473</u>	6.326	<u>76.6</u>
NI-WL	0.488	6.942	<u>76.0</u>
NI-WL-RD	<u>0.407</u>	<u>5.310</u>	74.1
NI-WL-RM	<u>0.432</u>	5.912	73.8
NI-WL-RM-KD	0.636	<u>5.181</u>	<u>76.5</u>

Bold underline, italic underline, and underline font highlights the first place, second place, and third place with same BufferSize, respectively.

^aP-Depth is short for Penetration Depth, P-Volume is short for Penetration Volume.

^bNI-WL, NI-WL-RD, NI-WL-RM, and NI-WL-RM-KD are four variants of our proposed method.

and cognition outside the brain.²⁷ Finally, the development of realistic test environments that specifically address CL capabilities of anthropomorphic grasping is another crucial factor for the advancement of continual grasping technology and need to be explored further.

Limitations of the study

Our approach only utilizes simulated data to improve the grasp performance and it is validated on dataset and in a simulator. Our demonstration here can be the first step toward a combination of neuro-inspired CL and anthropomorphic grasping. In the future, it not only needs to validate our approach on real robotics but also needs to find more sophisticated and effective methodologies that enable performance improvement using the data collected from both real task executions and simulators.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [RESOURCE AVAILABILITY](#)
 - Lead contact
 - Materials availability

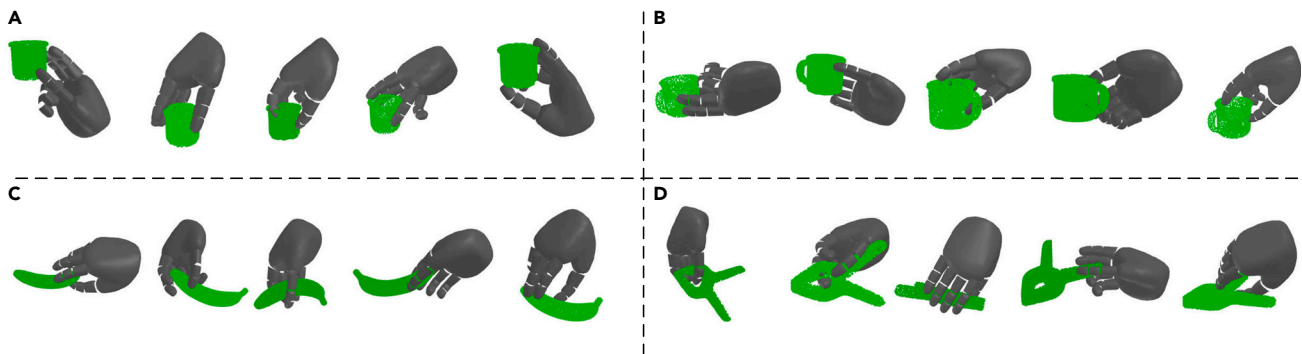


Figure 8. Qualitative grasps generated by NI-WL-RM-KD with 5K memory buffer on 4 objects from YCB object set

- (A) Grasps for object ycb 065-f cups scaled.
 (B) Grasps for object ycb 025.
 (C) Grasps for object ycb 011 banana scaled.
 (D) Grasps for object ycb 052 extra large clamp scaled.

- Data and code availability
- **METHOD DETAILS**
- Anthropomorphic robotic hand
- The used grasp model
- Description of compared methods
- Dataset
- Steps for the simulated experiments

ACKNOWLEDGMENTS

This work was supported by 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..

AUTHOR CONTRIBUTIONS

Conceptualization, W.L. and P.W.; Methodology & Investigation, W.L. and W.W.; Methodology & Software, W.L. and W.W.; Formal Analysis, W.L. and P.W.; Resources, P.W. and W.L.; Data Curation, W.W. and W.L.; Writing – Original Draft, W.L.; Writing – Review & Editing, W.L., W.W., and P.W.; Project Administration, P.W.; Supervision, P.W.

DECLARATION OF INTERESTS

The authors declare no competing interests.

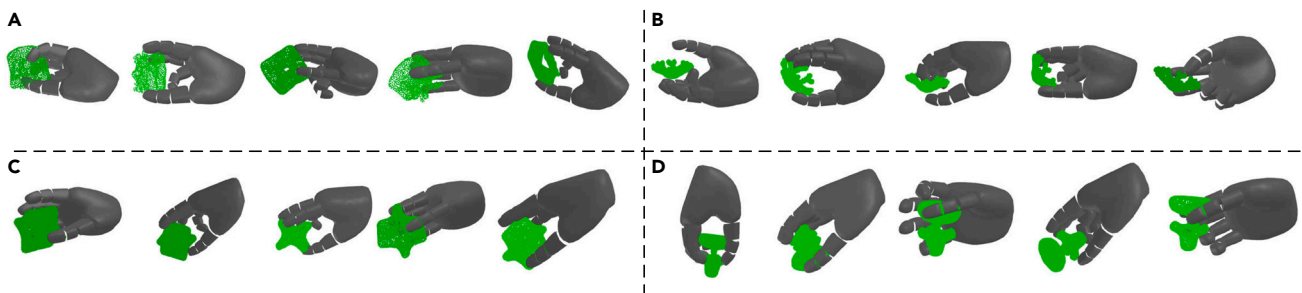


Figure 9. Qualitative grasps generated by NI-WL-RM-KD with 5K memory buffer on 4 objects from EGAD! object set

- (A) Grasps for object G5.
 (B) Grasps for object G6.
 (C) Grasps for object E1.
 (D) Grasps for object E4.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in their field of research or within their geographical location.

Received: January 5, 2023

Revised: March 13, 2023

Accepted: April 20, 2023

Published: April 25, 2023

REFERENCES

- McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–457.
- Hayes, T.L., Krishnan, G.P., Bazhenov, M., Siegelmann, H.T., Sejnowski, T.J., and Kanan, C. (2021). Replay in deep learning: current approaches and missing biological elements. *Neural Comput.* 33, 2908–2950.
- Abraham, W.C. (2008). Metaplasticity: tuning synapses and networks for plasticity. *Nat. Rev. Neurosci.* 9, 387.
- Schapiro, A.C., McDevitt, E.A., Rogers, T.T., Mednick, S.C., and Norman, K.A. (2018). Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nat. Commun.* 9, 3920–4011.
- Parisi, G.I., Kemker, R., Part, J.L., Kanan, C., and Wermter, S. (2019). Continual lifelong learning with neural networks: a review. *Neural Netw.* 113, 54–71.
- Tononi, G., and Cirelli, C. (2014). Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. *Neuron* 81, 12–34.
- Bao, C., Kim, T.-H., Hassanpoor Kalhori, A., and Kim, W.S. (2022). A 3d-printed neuromorphic humanoid hand for grasping unknown objects. *iScience* 25, 104119.
- Billard, A., and Kragic, D. (2019). Trends and challenges in robot manipulation. *Science* 364, eaat8414.
- Zhu, T., Wu, R., Lin, X., and Sun, Y. (2021). Toward human-like grasp: dexterous grasping via semantic representation of object-hand. In 2021 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 15721–15731. <https://doi.org/10.1109/ICCV48922.2021.01545>.
- Wei, W., Li, D., Wang, P., Li, Y., Li, W., Luo, Y., and Zhong, J. (2022). Dvvg: Deep Variational Grasp Generation for Dexterous Manipulation (IEEE Robotics and Automation Letters).
- Ficuciello, F., Migliozzi, A., Laudante, G., Falco, P., and Siciliano, B. (2019). Vision-based grasp learning of an anthropomorphic hand-arm system in a synergy-based control framework. *Sci. Robot.* 4, eaao4900.
- Kumaran, D., Hassabis, D., and McClelland, J.L. (2016). What learning systems do intelligent agents need? complementary learning systems theory updated. *Trends Cogn. Sci.* 20, 512–534.
- Sun, S., Calandriello, D., Hu, H., Li, A., and Titsias, M. (2022). Information-theoretic online memory selection for continual learning. In International Conference on Learning Representations.
- Mousavian, A., Eppner, C., and Fox, D. (2019). 6-dof graspnet: variational grasp generation for object manipulation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 2901–2910.
- Hasson, Y., Varol, G., Tzionas, D., Kalevatykh, I., Black, M.J., Laptev, I., and Schmid, C. (2019). Learning joint reconstruction of hands and manipulated objects. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 11807–11816.
- Jiang, H., Liu, S., Wang, J., and Wang, X. (2021). Hand-object contact consistency reasoning for human grasps generation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 11107–11116.
- Lomonaco, V., Pellegrini, L., Cossu, A., Carta, A., Graffieti, G., Hayes, T.L., De Lange, M., Masana, M., Pomponi, J., Van de Ven, G.M., et al. (2021). Avalanche: an end-to-end library for continual learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3600–3610.
- Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A.A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., et al. (2017). Overcoming catastrophic forgetting in neural networks. *Proc. Natl. Acad. Sci. USA* 114, 3521–3526.
- Zenke, F., Poole, B., and Ganguli, S. (2017). Continual learning through synaptic intelligence. In International Conference on Machine Learning (PMLR), pp. 3987–3995.
- Chaudhry, A., Rohrbach, M., Elhoseiny, M., Ajanthan, T., Dokania, P.K., Torr, P.H., and Ranzato, M. (2019). On tiny episodic memories in continual learning. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1902.10486>.
- Bang, J., Kim, H., Yoo, Y., Ha, J.-W., and Choi, J. (2021). Rainbow memory: continual learning with a memory of diverse samples. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8218–8227.
- Todorov, E., Erez, T., and Tassa, Y. (2012). Mujoco: a physics engine for model-based control. In IEEE/RSJ international conference on intelligent robots and systems (IEEE), pp. 5026–5033.
- Calli, B., Walsman, A., Singh, A., Srinivasa, S., Abbeel, P., and Dollar, A.M. (2015). Benchmarking in manipulation research: using the yale-cmu-berkeley object and model set. *IEEE Robot. Autom. Mag.* 22, 36–52.
- Morrison, D., Corke, P., and Leitner, J. (2020). Egd! an evolved grasping analysis dataset for diversity and reproducibility in robotic manipulation. *IEEE Robot. Autom. Lett.* 5, 4368–4375.
- Zou, X., Kolouri, S., Pilly, P.K., and Krichmar, J.L. (2020). Neuromodulated attention and goal-driven perception in uncertain domains. *Neural Netw.* 125, 56–69.
- Levinson, M., Kolenda, J.P., Alexandrou, G.J., Escanilla, O., Cleland, T.A., Smith, D.M., and Linster, C. (2020). Context-dependent odor learning requires the anterior olfactory nucleus. *Behav. Neurosci.* 134, 332–343.
- Emmons-Bell, M., Durant, F., Tung, A., Pietak, A., Miller, K., Kane, A., Martyniuk, C.J., Davidian, D., Morokuma, J., and Levin, M. (2019). Regenerative adaptation to electrochemical perturbation in planaria: a molecular analysis of physiological plasticity. *iScience* 22, 147–165.
- Liu, H., Wu, K., Meusel, P., Seitz, N., Hirzinger, G., Jin, M., Liu, Y., Fan, S., Lan, T., and Chen, Z. (2008). Multisensory five-finger dexterous hand: the dlr/hit hand ii. In IEEE/RSJ international conference on intelligent robots and systems (IEEE), pp. 3692–3697.
- Vitter, J.S. (1985). Random sampling with a reservoir. *ACM Trans. Math. Softw.* 11, 37–57.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
YCB dataset	YCBBenchmarks	https://www.ycbbenchmarks.com/object-set/
EGAD!	Dougsm	https://dougsm.github.io/egad/
Software and algorithms		
PyTorch	The Linux Foundation https://github.com/pytorch/pytorch	RRID:SCR_018536
Avalanche	ContinualAI	https://github.com/ContinualAI/avalanche
MuJoCo	DeepMind	https://github.com/deepmind/mujoco
Original code	This paper	https://github.com/WanyiLi/NICAG

RESOURCE AVAILABILITY

Lead contact

Further information and any related requests should be directed to and will be fulfilled by the lead contact, Peng Wang (peng wang@ia.ac.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data reported in this paper will be shared by the [lead contact](#) upon request.
- The code is deposited at the GitHub repository, <https://github.com/WanyiLi/NICAG>, and is publicly available as of the date of publication. A link to code has been included in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

METHOD DETAILS

Anthropomorphic robotic hand

The used anthropomorphic robotic hand is DLR/HIT Hand II.²⁸ It has five modular fingers with four joint and three active degrees of freedoms. We use a vector with 20 component $\theta \in \mathbb{R}^{20}$ to denote the joint configuration of the hand. The wrist pose \mathbf{p} of DLR/HIT Hand II is given in special Euclidean group $SE(3)$, 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 used grasp model

For grasp model M_g 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 DVG¹⁰ as a case study. We ignore the two auxiliary steps including object point completion and iterative grasp refinement for clarity.

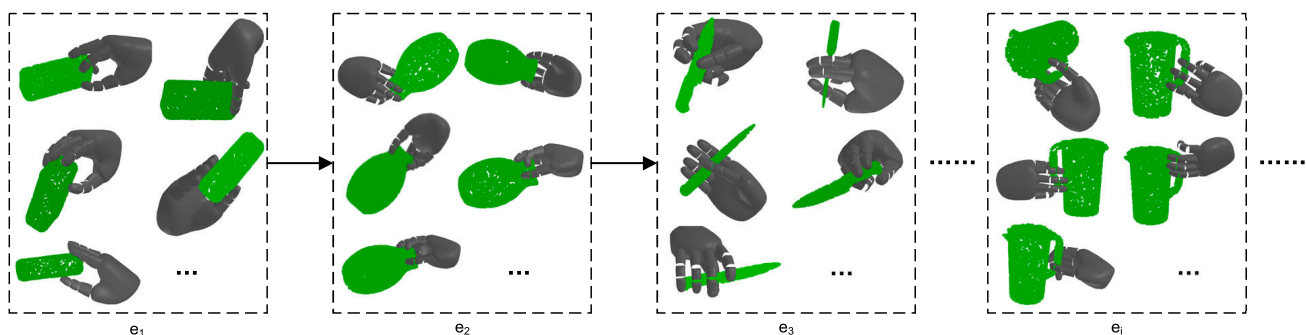
Description of compared methods

We compare the proposed approach with four typical continual learning approaches including elastic weight consolidation (EWC),¹⁸ synaptic intelligence (SI),¹⁹ experience replay (ER)²⁰ and rainbow memory (ER-RM),²¹ and two baselines, namely, Finetune and Ild-Offline. EWC¹⁸ and SI¹⁹ overcome forgetting with importance-based regularization. ER²⁰ is a simple but effective replay-based approach, which applies reservoir sampling²⁹ for memory update and random sampling for memory retrieval. Rainbow memory (ER-RM)²¹ enhances diversity of samples in a representative memory via a novel memory management strategy based on uncertainty and data augmentation. Diverse samples in memory are replayed to overcome

forgetting. Finetune incrementally finetunes the model without employing any continual learning strategy. Finetune can not overcome catastrophic forgetting and is considered as the naive baseline. Ild-Offline uses all the samples in the dataset in an offline manner to train the model, and is regarded as the oracle baseline.

Dataset

To evaluate our proposed framework and methods for continual learning of anthropomorphic grasping, we construct a sequential anthropomorphic grasping dataset based on Wei et al. . 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 set of grasp samples from an individual object, as shown in the below image. 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.



Grasp samples of each object in dataset as an experience

Steps for the simulated experiments

The simulated experiments are conducted in the physics-based simulator MuJoCo.²² 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 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 [experimental results](#).