



# Lifelong Robot Learning

---

终身学习综述

# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning

- 发表于： NIPS 2023 By Sony AI & 清华
- 目标：建立一个标准的终身学习测试平台
- 动机：机器人终身学习领域目前仍有大片空白
- 做了什么：测试了目前(2023)的终身学习方法和不同的Vision Encoder对效果的影响

# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning

- **探索五个关键研究课题：** 该基准的设计明确旨在推动对以下五个核心问题的研究：
  1. 如何高效地迁移陈述性知识、程序性知识或两者的混合。
  2. 如何为终身学习设计有效的策略网络架构。
  3. 如何为终身学习设计有效的算法。
  4. 终身学习器对于不同任务顺序的稳健性如何。
  5. 模型预训练对下游终身学习任务有何影响。
- **提供可扩展的任务和高质量数据：** 作者开发了一个可扩展的程序化任务生成流程，理论上可以创造无限多的任务。同时，他们创建了包含130个任务的四个任务套件，并为所有任务提供了高质量的人类遥操作演示数据，以支持高效的学习。

# 为什么区分陈述知识和程序知识？

- 运用陈述性知识来理解任务：

首先，它必须识别出场景中哪个是“番茄酱”，哪个是“盘子”，哪个是“篮子”。（*陈述性知识*）

它需要理解这些物体当前的位置关系，以及目标状态下“旁边”这个空间概念。（*陈述性知识*）

- 运用程序性知识来执行任务：

它需要知道如何移动机械臂到番茄酱的位置。（*程序性知识*）

它需要知道如何以正确的方式抓住番茄酱瓶。（*程序性知识*）

它需要知道如何将番茄酱平稳地移动到目标位置并放下。（*程序性知识*）

- 为什么区分陈述知识和程序知识？

- 论文作者认为，当一个终身学习的机器人失败时，我们很难判断问题出在哪里。

- 它是忘记了“碗”长什么样，或者忘记了“碗”放在哪里吗？（*陈述性知识的遗忘*）

- 还是它忘记了“抓取”这个动作该怎么做？（*程序性知识的遗忘*）

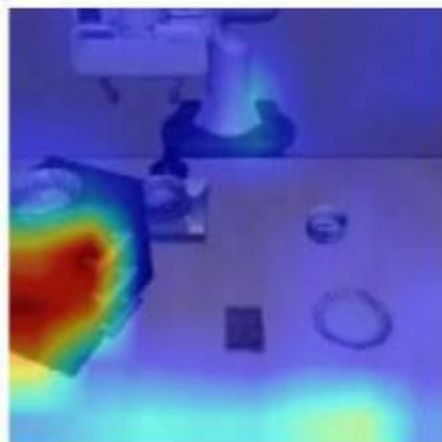
# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning

- **探索五个关键研究课题：** 该基准的设计明确旨在推动对以下五个核心问题的研究：
  1. 如何高效地迁移陈述性知识、程序性知识或两者的混合。
  2. 如何为终身学习设计有效的策略网络架构。
  3. 如何为终身学习设计有效的算法。
  4. 终身学习器对于不同任务顺序的稳健性如何。
  5. 模型预训练对下游终身学习任务有何影响。
- **提供可扩展的任务和高质量数据：** 作者开发了一个可扩展的程序化任务生成流程，理论上可以创造无限多的任务。同时，他们创建了包含130个任务的四个任务套件，并为所有任务提供了高质量的人类遥操作演示数据，以支持高效的学习。

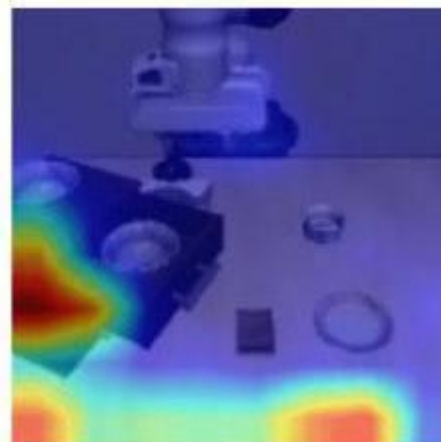
# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: 实验设计

- **LIBERO-SPATIAL:** 任务的目标和物体都相同，但物体的空间布局（位置关系）不同。这主要测试对 **陈述性知识** (空间信息) 的学习和记忆能力。
- **LIBERO-OBJECT:** 任务的目标和空间布局相似，但操作的物体不同。这主要测试对 **陈述性知识**（物体概念）的泛化能力。
- **LIBERO-GOAL:** 物体和布局都相同，但任务目标（需要执行的动作）不同。这主要测试对 **程序性知识**（“怎么做”）的学习能力。
- **LIBERO-100:** 一个包含100个任务的大型、混合套件，其中有混合的知识类型，更接近真实世界的复杂场景。

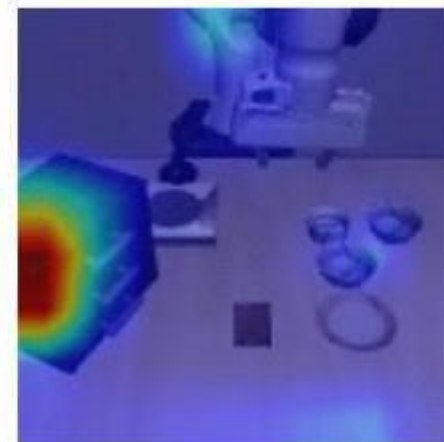
## LIBERO-SPATIAL



pick up the black bowl on the wooden cabinet and place it on the plate

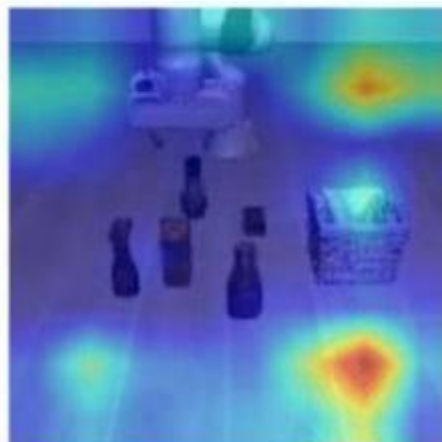


pick up the black bowl in the top drawer of the cabinet and place it on the plate

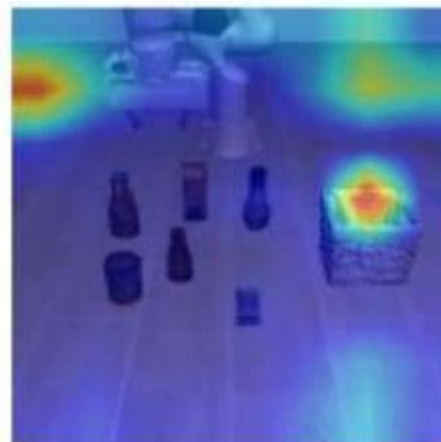


pick up the black bowl between the plate and the ramekin and place it on the plate

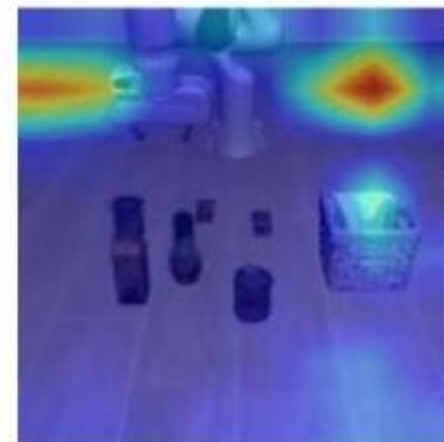
## LIBERO-OBJECT



pick up the orange juice and place it in the basket



pick up the ketchup and place it in the basket



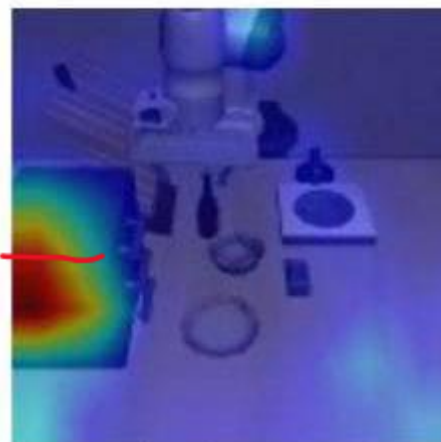
pick up the alphabet soup and place it in the basket



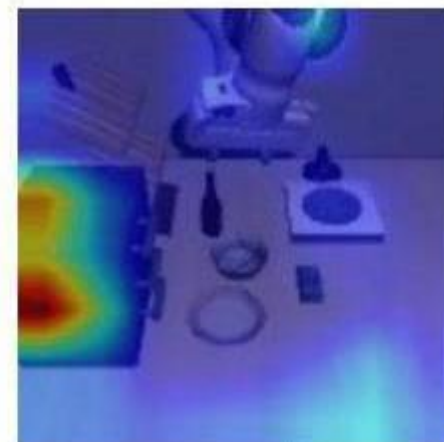
## LIBERO-GOAL



put the wine bottle  
on the rack

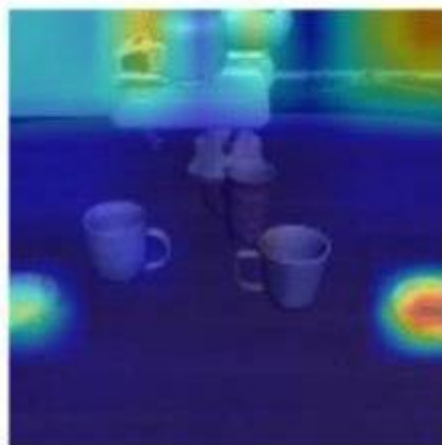


put the bowl on top  
of the cabinet



open the middle  
drawer of the  
cabinet

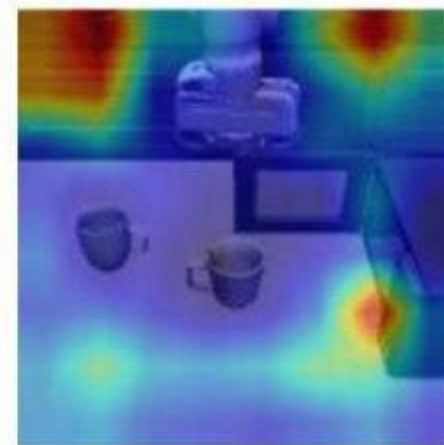
## LIBERO-LONG



put the white mug on  
the left plate and  
put the yellow and  
white mug on the  
right plate



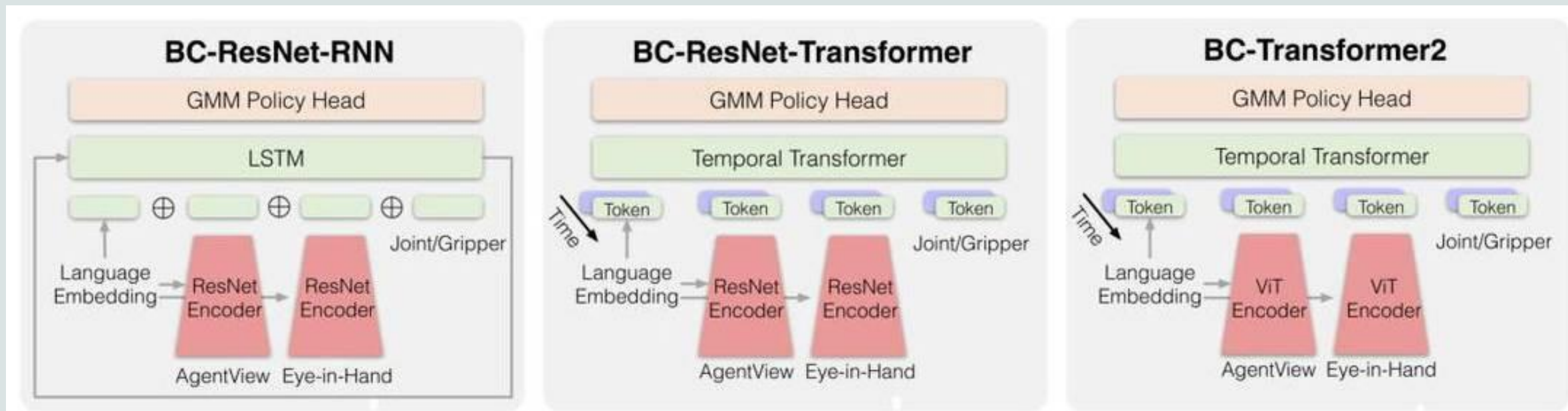
put both the  
alphabet soup and  
the tomato sauce in  
the basket



put the yellow and  
white mug in the  
microwave and close  
it



# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: BC Encoder 比较

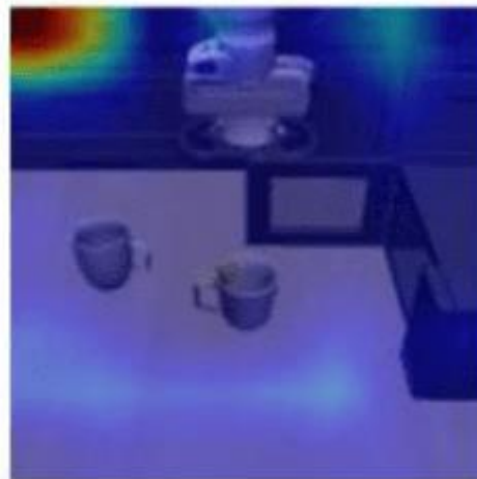


# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: BC Encoder 比较

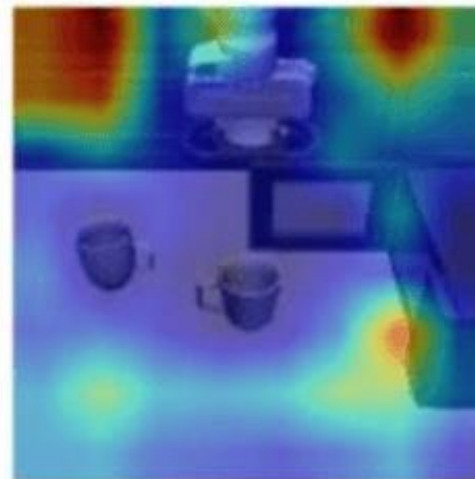
Policy Arch.	ER			PACKNET		
	FWT( $\uparrow$ )	NBT( $\downarrow$ )	AUC( $\uparrow$ )	FWT( $\uparrow$ )	NBT( $\downarrow$ )	AUC( $\uparrow$ )
LIBERO-LONG						
RESNET-RNN	$0.16 \pm 0.02$	<b><math>0.16 \pm 0.02</math></b>	$0.08 \pm 0.01$	$0.13 \pm 0.00$	$0.21 \pm 0.01$	$0.03 \pm 0.00$
RESNET-T	<b><math>0.48 \pm 0.02</math></b>	$0.32 \pm 0.04$	<b><math>0.32 \pm 0.01</math></b>	$0.22 \pm 0.01$	<b><math>0.08 \pm 0.01</math></b>	$0.25 \pm 0.00$
ViT-T	$0.38 \pm 0.05$	$0.29 \pm 0.06$	$0.25 \pm 0.02$	<b><math>0.36 \pm 0.01</math></b>	$0.14 \pm 0.01$	<b><math>0.34 \pm 0.01</math></b>
LIBERO-SPATIAL						
RESNET-RNN	$0.40 \pm 0.02$	$0.29 \pm 0.02$	$0.29 \pm 0.01$	$0.27 \pm 0.03$	$0.38 \pm 0.03$	$0.06 \pm 0.01$
RESNET-T	<b><math>0.65 \pm 0.03</math></b>	<b><math>0.27 \pm 0.03</math></b>	<b><math>0.56 \pm 0.01</math></b>	$0.55 \pm 0.01$	<b><math>0.07 \pm 0.02</math></b>	<b><math>0.63 \pm 0.00</math></b>
ViT-T	$0.63 \pm 0.01$	$0.29 \pm 0.02$	$0.50 \pm 0.02$	<b><math>0.57 \pm 0.04</math></b>	$0.15 \pm 0.00$	$0.59 \pm 0.03$
LIBERO-OBJECT						
RESNET-RNN	$0.30 \pm 0.01$	<b><math>0.27 \pm 0.05</math></b>	$0.17 \pm 0.05$	$0.29 \pm 0.02$	$0.35 \pm 0.02$	$0.13 \pm 0.01$
RESNET-T	$0.67 \pm 0.07$	$0.43 \pm 0.04$	$0.44 \pm 0.06$	<b><math>0.60 \pm 0.07</math></b>	<b><math>0.17 \pm 0.05</math></b>	<b><math>0.60 \pm 0.05</math></b>
ViT-T	<b><math>0.70 \pm 0.02</math></b>	$0.28 \pm 0.01$	<b><math>0.57 \pm 0.01</math></b>	$0.58 \pm 0.03$	$0.18 \pm 0.02$	$0.56 \pm 0.04$
LIBERO-GOAL						
RESNET-RNN	$0.41 \pm 0.00$	$0.35 \pm 0.01$	$0.26 \pm 0.01$	$0.32 \pm 0.03$	$0.37 \pm 0.04$	$0.11 \pm 0.01$
RESNET-T	<b><math>0.64 \pm 0.01</math></b>	<b><math>0.34 \pm 0.02</math></b>	<b><math>0.49 \pm 0.02</math></b>	$0.63 \pm 0.02$	<b><math>0.06 \pm 0.01</math></b>	$0.75 \pm 0.01$
ViT-T	$0.57 \pm 0.00$	$0.40 \pm 0.02$	$0.38 \pm 0.01$	<b><math>0.69 \pm 0.02</math></b>	$0.08 \pm 0.01$	<b><math>0.76 \pm 0.02</math></b>

# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: BC Encoder 比较

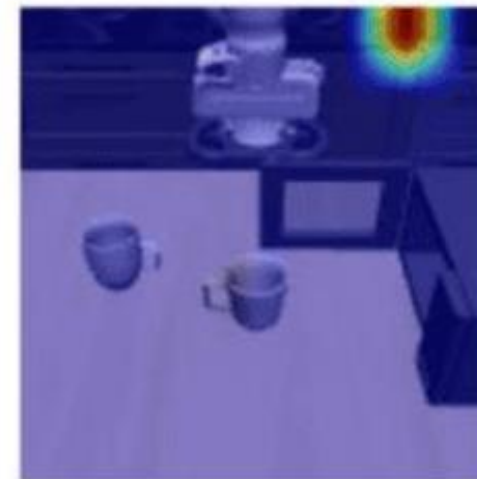
**Policy after  
training task 10  
on task 10**



**ResNet-RNN**



**ResNet-T**



**ViT-T**

图 6: 注意力可视化, 比较了不同架构 (ResNet-RNN、ResNet-T 和 ViT-T) 在任务执行期间的关注点。

这些可视化显示, 基于 Transformer 的架构会形成更集中的注意力模式, 目标是与任务相关的对象, 而基于 RNN 的方法则倾向于具有更分散的注意力。

# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: 预训练的影响

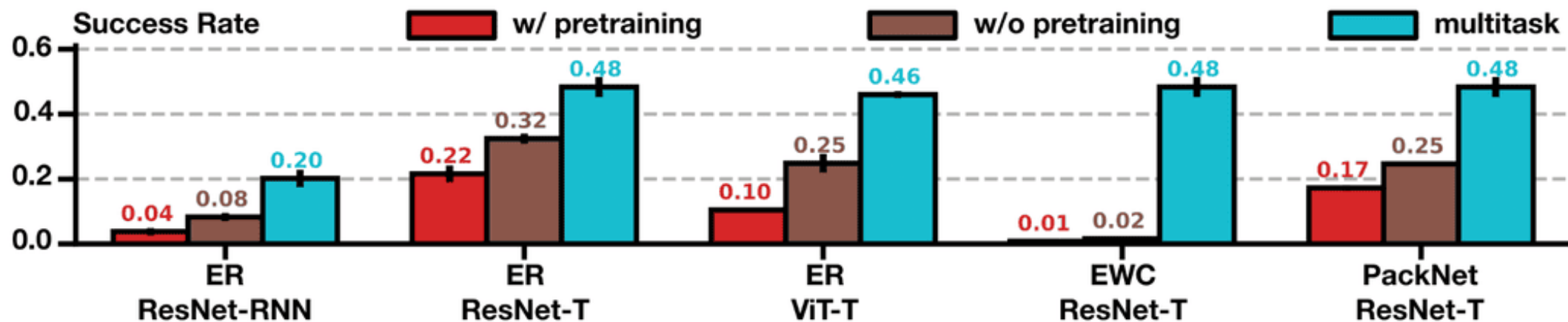


Figure 5: Performance of different combinations of algorithms and architectures without pretraining or with pretraining. The multi-task learning performance is also included for reference.

# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: 终身学习方法比较

- **ER (Experience Replay):** 定时复习。它会存储一小部分旧任务的数据，在学习新任务时拿出来一起“复习”，以对抗遗忘。
- **EWC (Elastic Weight Consolidation):** **Regularization-based** 的代表。它通过计算并限制重要参数的改动，来保护旧任务学到的知识不被新任务的学习过程破坏。
- **PACKNET:** 动态架构法 (**Dynamic Architecture-based**) 的代表。它为每个新任务分配一部分网络参数，并冻结旧任务的参数，从而在结构上避免遗忘。



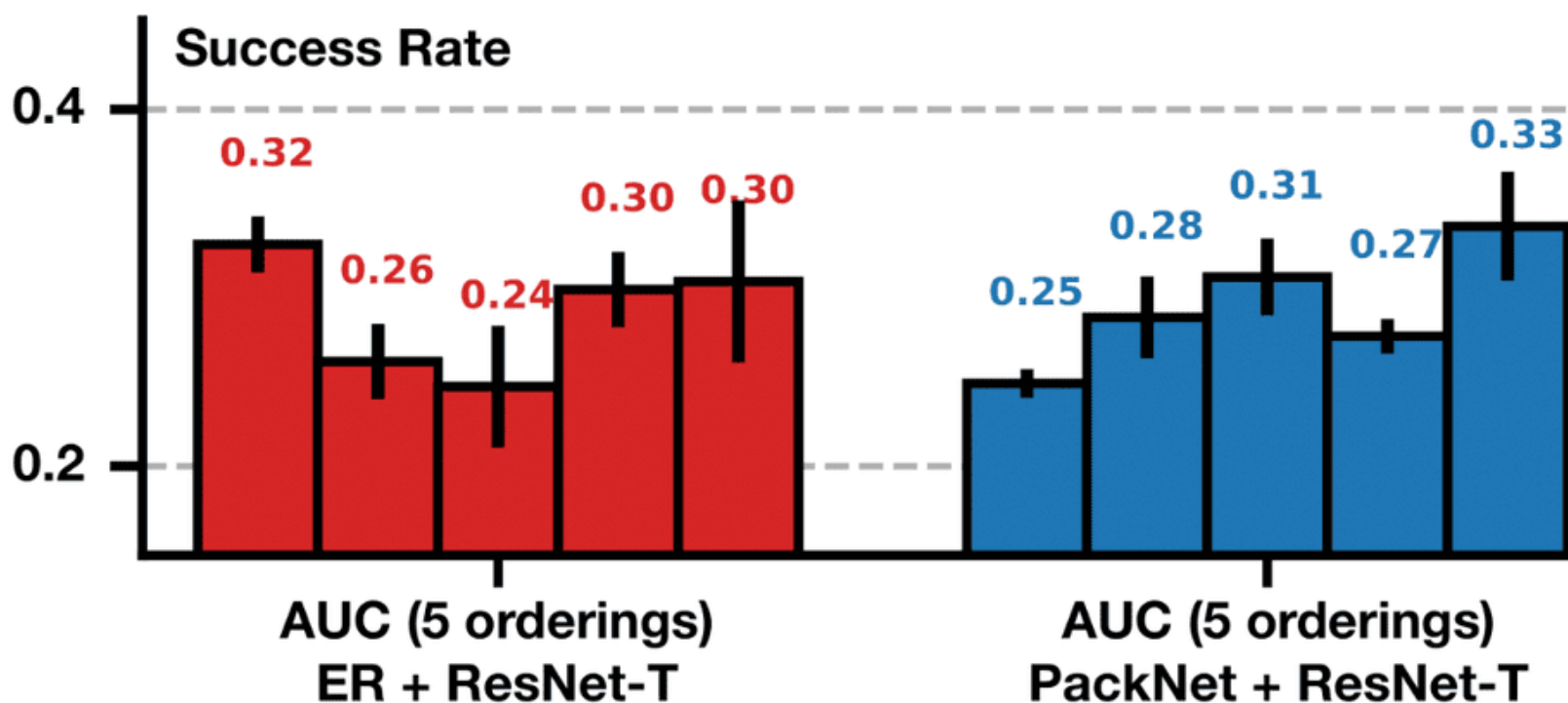
# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: 终身学习方法比较

Lifelong Algo.	FWT( $\uparrow$ )	NBT( $\downarrow$ )	AUC( $\uparrow$ )	FWT( $\uparrow$ )	NBT( $\downarrow$ )	AUC( $\uparrow$ )
	LIBERO-LONG			LIBERO-SPATIAL		
SEQL	<b>0.54</b> $\pm$ 0.01	0.63 $\pm$ 0.01	0.15 $\pm$ 0.00	<b>0.72</b> $\pm$ 0.01	0.81 $\pm$ 0.01	0.20 $\pm$ 0.01
ER	0.48 $\pm$ 0.02	0.32 $\pm$ 0.04	<b>0.32</b> $\pm$ 0.01	0.65 $\pm$ 0.03	0.27 $\pm$ 0.03	0.56 $\pm$ 0.01
EWC	0.13 $\pm$ 0.02	0.22 $\pm$ 0.03	0.02 $\pm$ 0.00	0.23 $\pm$ 0.01	0.33 $\pm$ 0.01	0.06 $\pm$ 0.01
PACKNET	0.22 $\pm$ 0.01	<b>0.08</b> $\pm$ 0.01	0.25 $\pm$ 0.00	0.55 $\pm$ 0.01	<b>0.07</b> $\pm$ 0.02	<b>0.63</b> $\pm$ 0.00
MTL			0.48 $\pm$ 0.01			0.83 $\pm$ 0.00
	LIBERO-OBJECT			LIBERO-GOAL		
SEQL	<b>0.78</b> $\pm$ 0.04	0.76 $\pm$ 0.04	0.26 $\pm$ 0.02	<b>0.77</b> $\pm$ 0.01	0.82 $\pm$ 0.01	0.22 $\pm$ 0.00
ER	0.67 $\pm$ 0.07	0.43 $\pm$ 0.04	0.44 $\pm$ 0.06	0.64 $\pm$ 0.01	0.34 $\pm$ 0.02	0.49 $\pm$ 0.02
EWC	0.56 $\pm$ 0.03	0.69 $\pm$ 0.02	0.16 $\pm$ 0.02	0.32 $\pm$ 0.02	0.48 $\pm$ 0.03	0.06 $\pm$ 0.00
PACKNET	0.60 $\pm$ 0.07	<b>0.17</b> $\pm$ 0.05	<b>0.60</b> $\pm$ 0.05	0.63 $\pm$ 0.02	<b>0.06</b> $\pm$ 0.01	<b>0.75</b> $\pm$ 0.01
MTL			0.54 $\pm$ 0.02			0.80 $\pm$ 0.01

Table 2: Performance of three lifelong algorithms and the SEQL and MTL baselines on the four task suites, where the policy is fixed to be RESNET-T. Results are averaged over three seeds and we report the mean and standard error. The best performance is **bolded**, and colored in **purple** if the improvement is statistically significant over other algorithms, when a two-tailed, Student’s t-test under equal sample sizes and unequal variance is applied with a  $p$ -value of 0.05.



# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: 任务顺序的影响



# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: Take Away

- 架构设计与算法设计同等重要，对于终身机器人学习而言，基于Transformer的架构通常优于基于RNN的方法。
- 不同的神经网络架构擅长不同类型的迁移，其中视觉Transformer在以物体为中心的任务中表现出优势，而卷积网络在空间推理任务中表现良好。
- 当前的终身学习算法面临着前向和后向迁移之间的权衡，没有单一算法在这两方面都表现出色。
- 任务顺序显著影响性能，突出了设计在终身学习环境中的重要性。
- 传统的预训练方法可能不利于机器人操作任务中的终身学习，这表明需要专门为增量学习设计的预训练方法。

# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: 环境和算力要求

- 模拟环境：Robosuite
- 算力要求：1 X A100 or 1 X A40
- 代码已在 **Github** 上开源
- 数据集也可以在项目网站上获取

# LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning: 我的想法

- LIBERO 没有测试 Diffusion Policy + Life Long Learning
- LIBERO 没有测试 Pretrain LLM 的影响
- LIBERO 中的成功率仍不算太高 (与其他 Few shot Learning 相比)
- 对比  $\pi_0$ ,  $\pi_0$  是通过冻结 VLM + Finetune Action Head 来做的 "Lifelong Learning"
- 我认为也可以在  $\pi_0$  的基础上尝试增加其他 Lifelong Learning 方法



Thanks for  
Listening!



# Appendix

