



机器人抓取与操作第一章作业提示

主讲人 快乐风男



- 第一部分：知识基础概述
- 第二部分：实践原理提示
- 第三部分：改进思路提升

- 这门课默认大家具备的软件工具基础

- Ubuntu 或者 Windows 使用Conda 配置python 开发环境

- Conda 官方文档

- <https://docs.conda.io/en/latest/>

- Docker的使用

- Docker 常用命令大全

- https://blog.csdn.net/m0_63653444/article/details/139130248

- 这门课需要大家具备的知识基础
 - 深度学习模型基础
 - Pytorch搭建模型
 - CVAE架构、transformer架构
 - 机器人学基础
 - 基本坐标变换、前向运动学、逆向运动学
 - 机器人运动规划算法

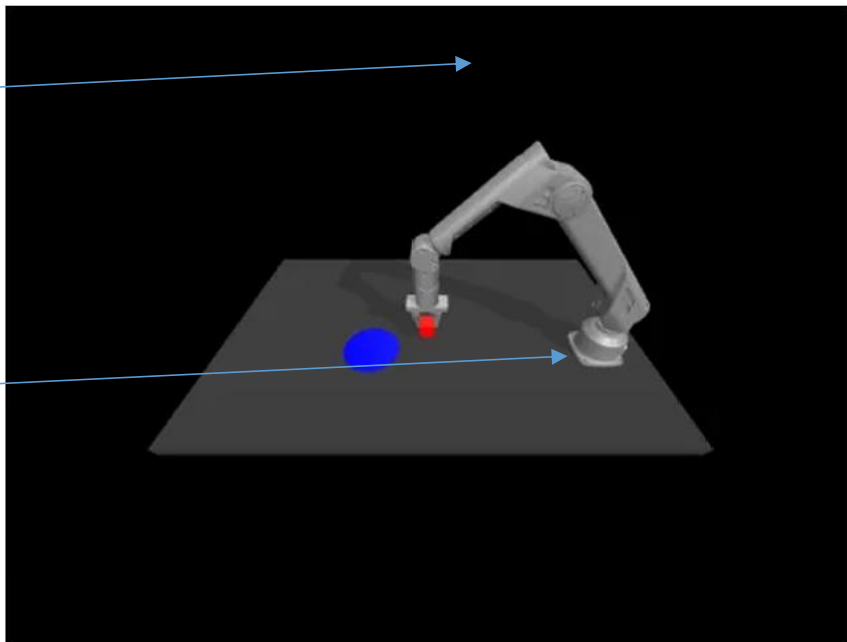
- 第一部分：知识基础概述
- **第二部分：实践原理提示**
- 第三部分：改进思路提示

实践原理提示

- 任务环境简介 使用ACT在mujoco仿真环境中训练GalaXea的机械臂A1，让它能够抓起红色方块并且成功放在蓝色圆盘上。

Top相机位置

6关节+夹
爪=7dof



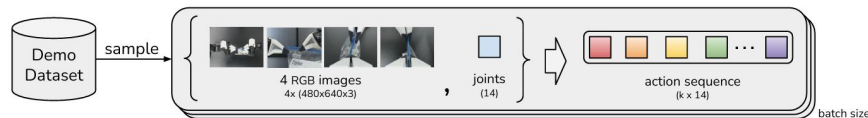
实践原理提示

● ACT简介：训练部分

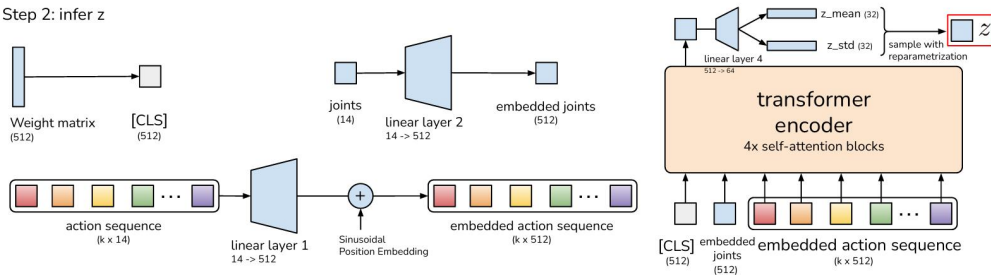
注意：

1. 模型输入不一样：本次任务相机只有一个，关节有7个
2. ACT详细代码建议参考policy.py中ACT高度复用DETR代码，如果想深入细节建议参考：
<https://github.com/facebookresearch/detr>
3. Style Variable Z是用来做正则化的关节输出的

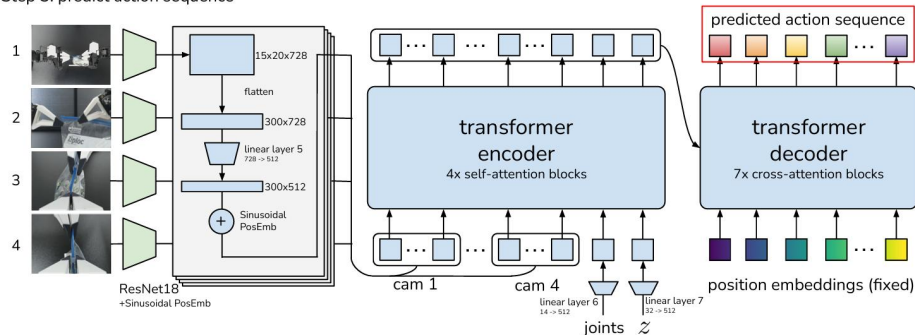
Step 1: sample data



Step 2: infer z



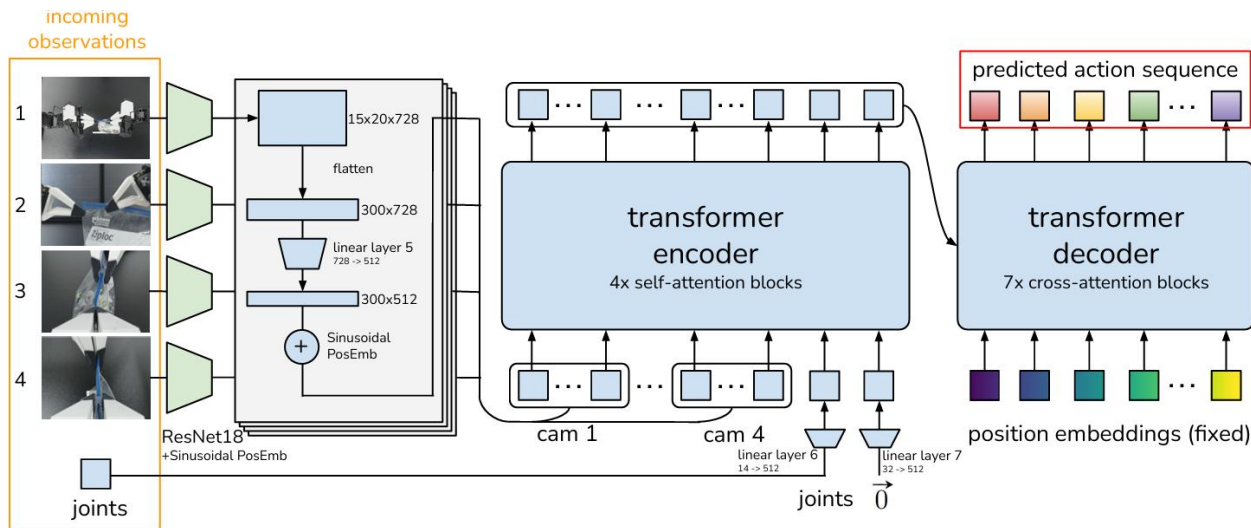
Step 3: predict action sequence



实践原理提示

●ACT简介：推理部分1

和传统的
CVAE有点区
别：推理阶
段把Style
Variable Z设置
为0



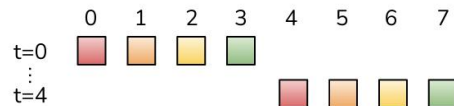
实践原理提示

●ACT简介：推理部分2

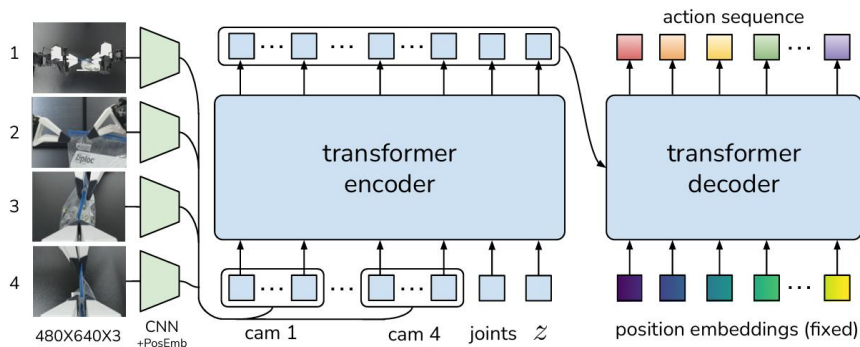
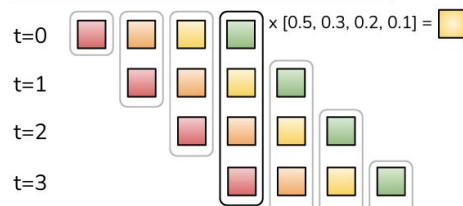
作者采用了至关重要的
ActionTrunking+Temporal Ensemble的
trick，即一次观测输入会连续推断
trunc_size个步长的控制输出，然后
连续观测的控制输出会被聚合在一起
(tempora ensemble) 平滑噪声。

这个trick是用来处理Compounding
Error的重要技术

Action Chunking



Action Chunking + Temporal Ensemble



- 第一部分：概述
- 第二部分：理论计算部分
- 第三部分：改进思路提示

改进思路提示

- 想到如何调优需要大家具备一定的深度学习基础
深度学习要优化好一个模型一般从以下方面入手：
 - 数据集（e. g. 让分布更均匀）
 - 模型（e. g. 增加复杂度或者架构更贴合任务）
 - 优化方式（e. g. 改进loss改进优化器）

●改进数据集的思路

生成数据的时候初始方块位置随机，脚本随机生成50个位置，注意脚本里面没有任何保证50个位置能够均匀分布，有的同学运气好生成的均匀，有的同学运气不好生成得位置训练和测试的分布差很远就导致了大家测试结果千差万别。所以一个有效的改进方式是直接生成一组起始方块位置分布均匀的数据集。

改进思路提示

●改进模型的思路

一般来说只要数据集的分布没有大问题，加大模型的复杂度都会涨点，这里给出一些方式做参考：

- ResNet的深度
- Transformer hidden_dim的大小和Head数量
- Chunk_Size 的增加

改进思路提示

●改进优化方式的思路

改进优化方式一般来说没有前两种方法直观，这里给出一些方式作参考：

- 调整Lreconst 与 Lreg的权重配比
- Lreconst尝试换成 L1loss或者Huber Loss
- 尝试调整优化器参数

感谢各位聆听 !
Thanks for Listening

