

扩散策略：通过行动扩散进行视觉运动策略学习

摘要

本研究探索了基于扩散模型的机器人视觉运动策略学习方法，旨在提升机器人在复杂环境中的操作能力。通过将条件去噪扩散模型应用于机器人的动作序列生成，我们展示了该方法在执行精确操作任务时的高效性和准确性。与传统的行为克隆和强化学习方法相比，扩散模型在训练稳定性和效率方面表现出显著优势。实验结果验证了该模型在提高机器人任务成功率和操作精度方面的潜力，预示着其在工业自动化和服务机器人领域的广泛应用前景。

关键词：具身智能；diffusion policy

1 引言

随着人工智能和机器人技术的快速发展，机器人在工业生产、日常服务和复杂环境任务中的应用越来越广泛。然而，传统的机器人行为生成方法，如直接的行为克隆或简单的监督学习，常常在复杂任务执行、多模态动作分布处理以及训练稳定性方面遇到限制。这些限制显著阻碍了机器人技术的进一步应用和发展，特别是在需要精细操作和高度自适应能力的场景中 [7, 18]。

为了克服这些挑战，本研究引入了一种基于扩散模型的视觉运动策略学习方法。扩散模型，作为一种先进的生成模型，已在图像和音频生成领域显示出其卓越的性能 [12]。最近，将其应用于机器人视觉运动策略学习，展示了处理复杂动作序列和多模态动作分布的潜力 [14, 19]。这种基于条件去噪扩散过程的方法不仅能迭代地精化噪声生成高质量的行动指令，还能通过学习动作分布的梯度优化机器人的行为输出 [18, 20]。

本研究的主要目标是通过实验验证扩散模型在多个机器人操作任务中的应用效果。将评估其在不同任务类型中的表现，包括精确搬运、复杂装配及敏感操作等，以此来测试其对于多模态动作分布的处理能力及在复杂序列任务中的应用潜力。此外，将扩散模型与当前最先进的机器人学习方法进行比较，以体现其在训练稳定性、效率和动作准确性方面的优势 [8]。

基于扩散策略的机器人视觉运动策略学习不仅具有重要的学术价值，也有望在工业自动化、服务机器人以及更广泛的智能系统中找到实际应用，具有重大的社会和经济意义。通过这项研究，期望能够提高机器人执行复杂任务的能力，优化其在实际工作环境中的表现，同时推动扩散模型在机器人学领域的进一步研究和应用 [7, 10]。

2 相关工作

2.1 行为克隆技术

行为克隆技术是通过观察专家的行为直接学习控制策略的一种方法，广泛应用于机器人学和自动驾驶车辆领域 [1]。这部分工作主要包括两种策略学习方法：

- **直接回归方法：**此类方法通常采用深度神经网络，如卷积神经网络（CNN）或多层感知机（MLP），以实现从状态到动作的映射 [1]。直接回归的简洁性虽然受到青睐，但其对训练数据的依赖性及泛化能力的局限性也不容忽视。
- **序列决策方法：**在动态环境中，行为序列的时间依赖性成为提升策略性能的关键。通过采用循环神经网络（RNN）或长短期记忆网络（LSTM），序列决策方法有效地捕捉了行为之间的时序关系，提升了对环境动态的适应能力 [2]。

2.2 强化学习方法

与行为克隆不同，强化学习通过与环境的直接交互来优化策略，是求解决策问题的一种更为动态的方法 [17]。强化学习通过奖励机制自主探索有效策略，不依赖预先定义的行为数据。

- **模型自由的 RL：**模型自由的强化学习方法，如 Q 学习和策略梯度方法，直接对策略或值函数进行优化 [11,17]。这些方法虽然能在复杂的决策环境中实现端到端的学习，但其样本效率较低，且在初期依赖大量的探索。
- **模型依赖的 RL：**模型依赖的强化学习通过构建环境模型来预测未来的状态和奖励，从而减少对实际环境交互的需求 [17]。这种方法能有效加速学习过程，但其性能高度依赖于模型的准确性。

2.3 生成模型在策略学习中的应用

生成模型，如生成对抗网络（GAN）和变分自编码器（VAE），在策略学习中展示了处理复杂决策环境和多模态动作分布的能力 [6,9]。这些技术通过生成高质量的数据，提供了一种有效的方法来增强策略学习过程，尤其在数据稀缺或需要强泛化的场景中。

- **使用 GAN 进行策略学习：**生成对抗网络（GAN）利用其生成器和判别器之间的对抗过程，能够产生逼真的动作数据 [6]。这些数据不仅质量高，而且能够捕捉到复杂的动作分布，使其成为训练策略网络的理想选择。GAN 特别适用于数据稀缺的情况，能够从有限的样本中学习并生成新的、未见过的数据实例。
- **使用 VAE 进行策略学习：**变分自编码器（VAE）通过其编码和解码的过程，学习数据的潜在表示，从而生成新的观察数据 [9]。这种方法不仅增强了策略网络的训练过程，还提高了模型在面对未知状态时的泛化能力。VAE 特别适合于处理需要强泛化能力的任务，因为它能够通过学习一个连续的潜在空间来揭示数据背后的关键因素。

2.4 扩散模型在图像领域的应用及其对机器人策略的影响

扩散模型作为一种先进的生成技术，最近在图像处理领域得到了广泛的应用 [7,12]。这些模型通过逐步去噪过程生成图像，不仅能够产生高质量的视觉内容，而且在机器人视觉系统的开发中显示出巨大的潜力 [15]。扩散模型能够帮助机器人更准确地解析和理解其操作环境，从而在实际任务中做出更合理的策略决策。

- **高质量图像生成：**扩散模型能够生成高分辨率和高细节的图像，这对于机器人进行环境识别和对象检测尤为重要 [12,15]。通过改善视觉输入的质量，机器人能够更准确地识别其周围的物体和障碍，从而优化其导航和操作策略。
- **条件图像合成：**在任务指导的条件下，扩散模型可以生成与特定指令或描述相匹配的图像 [16]，这为机器人提供了执行复杂视觉任务的能力，如在复杂环境中寻找特定物体或人员。这种能力极大地增强了机器人在救援、监视和服务行业中的应用潜力。
- **图像编辑与恢复：**扩散模型在图像恢复方面的应用，可以使机器人在视觉受损或不完整的情况下，恢复出有用的视觉信息 [12]。这对于在极端或复杂的环境中操作的机器人尤其重要，例如在烟雾或尘埃中导航。

3 本文方法

本文 [3] 提出了一种基于扩散过程的机器人视觉运动控制策略，通过分析不同的网络架构，展示了该策略在处理复杂动作序列中的有效性。如图 1 所示，该策略包括以下几个关键组成部分：

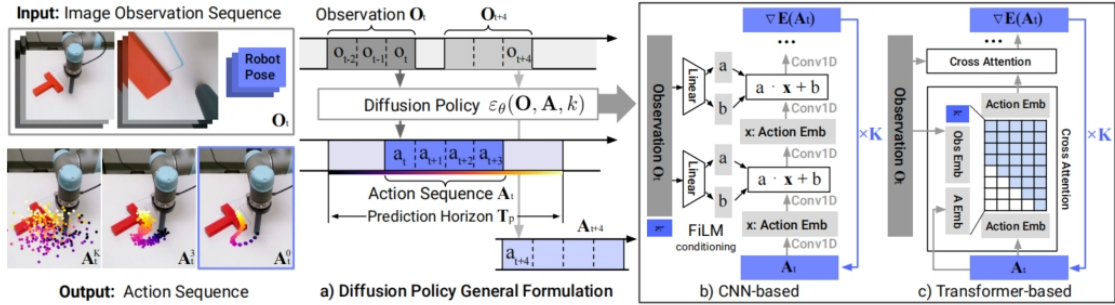


图 1. 方法示意图

- **输入与观测序列：**策略接受一系列图像观测作为输入，这些图像捕捉了机器人及其操作环境的实时状态 [5]。
- **扩散策略总体框架（General Formulation）：**在每一个时间步骤 t ，策略使用最新的观测数据 O_t 来预测一系列动作 A_t 。这一过程涉及到多个时间步的动作预测，以保证动作的连贯性和适时性 [7]。
- **基于 CNN 的方法：**采用卷积神经网络（CNN）和特征级线性调制（FiLM）技术对观测特征进行处理，以增强模型对关键视觉信息的响应能力。每个卷积层都会根据观测数据进行调整，以优化后续的动作预测 [13]。

- **基于 Transformer 的方法：**使用 Transformer 架构来处理高频动态变化，通过在多头自注意力机制中融入观测数据，精确调控每一步的动作输出。这种方法特别适合处理需要快速响应的复杂动作序列 [4]。
- **输出动作序列：**最终输出的动作序列 A_t 通过多次迭代的去噪过程得到精细化，确保动作的准确性和适应性 [18]。

3.1 扩散策略概述

本文介绍了一种新的机器人行为生成方法——扩散策略，将机器人的视觉运动策略表述为条件降噪扩散过程。通过在多个机器人操控基准测试中的对比分析，发现该策略在各项任务中均显著优于现有的最先进方法，平均改进率达到 46.9%。

3.1.1 扩散策略的基本原理

扩散策略源自于降噪扩散概率模型（DDPMs），该模型是一类生成模型，通过模拟去噪过程逐步生成输出。在机器人视觉运动控制的应用中，扩散策略通过迭代地优化动作分布的梯度场，并在推断过程中采用随机朗之万动力学步骤，从而能够表达复杂的多模态动作分布，适用于高维动作空间，并展示出令人印象深刻的训练稳定性，具体实现逻辑如下伪代码所示：

Algorithm 1 Diffusion Strategy for Robotic Motion Control

```

1: Input: Initial state  $x_0$ , number of diffusion steps  $N$ 
2: Output: Generated action sequence  $A$ 
3: Initialize  $x = x_0$ 
4: for  $t = 1$  to  $N$  do
5:   Compute noise level  $\sigma_t$  using a predefined schedule
6:   Sample noise  $\epsilon \sim N(0, \sigma_t^2)$ 
7:    $x = x + \epsilon$  ▷ Forward diffusion process (adding noise)
8: end for
9: Initialize latent state  $z = x$ 
10: for  $t = N$  down to 1 do
11:   Estimate gradient  $\nabla_x L(x, t)$  where  $L$  is the loss based on the action distribution
12:   Compute reverse noise level  $\beta_t$  using a predefined schedule
13:   Sample reverse noise  $\eta \sim N(0, \beta_t^2)$ 
14:    $z = z - \nabla_x L(z, t) \cdot dt + \eta$  ▷ Reverse diffusion process (denoising step)
15: end for
16: Decode the final latent state  $z$  to generate the action sequence  $A$ 
17: return  $A$ 

```

3.1.2 扩散策略的数学模型

在扩散策略中，从高斯噪声开始，模型执行 K 次去噪迭代，生成一系列噪声逐渐减少的中间动作，直至得到无噪声的最终输出。该过程可以用以下方程表示：

$$x_{k-1} = \alpha_k x_k - \frac{\beta_k}{\sqrt{1 - \alpha_k^2}} \epsilon_\theta(x_k, k) + \sigma_k \zeta_k$$

其中， x_k 为第 k 步的动作表示， ϵ_θ 为噪声预测网络， α_k ， β_k ，和 σ_k 是迭代步骤中的调节参数， ζ_k 是加入的高斯噪声，用于确保过程的随机性。

通过这种方法，扩散策略能够在物理机器人上学习视觉运动策略，充分发挥扩散模型在生成模型领域的强大能力，为新一代政策学习技术提供动力。

3.2 网络架构

扩散策略在网络架构选择上进行了创新，提出了两种主要的网络架构：卷积神经网络（CNN）和 Transformers。CNN 用于处理具有时间连续性的行动序列，而 Transformers 则用于处理需要高频动态变化的任务。

3.2.1 卷积神经网络（CNN）

在本研究中，卷积神经网络（CNN）被用于处理机器人感知系统连续捕获的图像数据。利用其深层卷积架构，CNN 能够从这些图像中提取出高度复杂的特征，这些特征对于后续的动作预测过程至关重要。卷积层通过不同的滤波器学习图像的多层次特征，从基本的纹理和形状到复杂的对象组合，每个滤波器自动捕捉到从简单到复杂的视觉信息。这种层次化的信息提取方法为动作决策提供了必要的输入，尤其在需要从动态环境中迅速做出反应的场景中表现出极高的效率。

此外，CNN 不仅用于特征提取，还通过与特征级线性调制（FiLM）技术的结合，增强了模型对特定任务条件的适应性。FiLM 层通过动态调整卷积层输出的特征图的缩放和偏移，根据当前的环境和任务需求优化特征的表达，从而提高整体网络的表现和灵活性。这一调整机制是通过学习如何最佳地整合额外的任务相关信息，例如当前的观测 O_t 和预定的动作 A_t ，以实现最优的动作输出。

3.2.2 Transformers

模型介绍

在本研究中，采用了基于 Transformer 的架构来处理时间序列数据的扩散问题。Transformer 架构通过利用自注意力机制：

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

其中， Q, K, V 分别代表查询（Query），键（Key），和值（Value）矩阵， d_k 是键的维度，这个机制允许模型在每个时间步骤对全局信息进行加权，从而提取出最关键的特征。

利用该机制，捕获长期依赖关系。这一机制使得每个输入元素能全面地反映整个输入序列的信息，极大增强了模型对时间序列变化的敏感度和响应速度，特别适合于机器人视觉运

Algorithm 2 Time-Series Data Convolutional Network Components

```
1: class Downsample1d:
2:     def __init__(self, dim):
3:         self.conv = Conv1d(dim, dim, 3, stride=2, padding=1)
4:     def forward(self, x):
5:         return self.conv(x)
6:
7: class Upsample1d:
8:     def __init__(self, dim):
9:         self.conv = ConvTranspose1d(dim, dim, 4, stride=2, padding=1)
10:    def forward(self, x):
11:        return self.conv(x)
12:
13: class Conv1dBlock:
14:     def __init__(self, inp_channels, out_channels, kernel_size, n_groups=8):
15:         self.block = Sequential(
16:             Conv1d(inp_channels, out_channels, kernel_size, padding=kernel_size // 2),
17:             GroupNorm(n_groups, out_channels),
18:             Mish())
19:     def forward(self, x):
20:         return self.block(x)
21:
22: function test():
23:     cb = Conv1dBlock(256, 128, kernel_size=3)
24:     x = zeros(1, 256, 16)
25:     o = cb(x)
26:     print shape of o
```

动策略学习。有效地捕获了长距离的时间依赖关系，这对于理解和预测在复杂动态环境中的连续行动序列至关重要。其核心优势在于能够并行处理整个序列，与传统的逐步处理方式相比，显著提高了计算效率和响应速度。

高频动态处理

Transformers 通过其多头注意力机制：

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O$$

$$\text{where head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$$

其中， W^O, W_i^Q, W_i^K, W_i^V 是可学习的参数矩阵， h 是注意力头的数量。

这种结构不仅增强了模型的学习能力，也提高了对未来状态的预测准确性，允许模型在多个维度并行处理信息，快速地调整对环境变化的响应策略。这不仅提升了处理效率，同时通过精准地调节注意力分配，最小化了过平滑现象，确保了在需精确控制速度的任务中能够达到优越的性能表现。在这些任务中，模型必须能够迅速识别并响应关键的环境变量变化。该技术通过分散注意力，能够在不同的表示子空间中同时评估多个环境因素，从而实现对复杂动态的高效处理。

3.3 ResNet-18 的改进

该研究对标准的 ResNet-18 架构进行了针对性的改进，以适应机器人视觉运动控制的复杂需求。主要改进集中在网络的池化和归一化技术上，旨在提升模型的特征提取能力和训练稳定性。

3.3.1 空间最大池化

该研究将 ResNet-18 中的标准全局平均池化层替换为空间最大池化层，以强化模型对关键视觉信息的捕捉能力：

- **操作定义：** 在每个卷积层后应用空间最大池化，通过选择特征图中每个通道的最大值来代替平均值。
- **数学表达式：** $\text{MaxPool}(x) = \max_{i,j}(x_{i,j})$ ，其中 $x_{i,j}$ 表示在位置 (i, j) 的像素值。
- **目的：** 最大池化更有效地保留了图像的显著特征，特别是在特征表示的局部区域中，这对于精确识别机器人环境中的关键元素至关重要。

3.3.2 群归一化

为了提高模型在不同批次大小训练时的稳定性，该研究引入了群归一化，代替了传统的批归一化：

- **操作定义：** 在每个卷积层后实施群归一化，通过计算属于同一组的特征通道的均值和方差来进行归一化。

- **数学表达式：** $y = \frac{x-\mu}{\sigma+\epsilon} \times \gamma + \beta$ ，其中 μ 和 σ 是组内均值和标准差， γ 和 β 是可学习的尺度和偏移参数。
- **目的：** 群归一化使模型在训练时更加稳定，尤其是在小批量数据或不同硬件资源下运行时，从而提高了模型在动态环境下的泛化能力和准确性。

这些改进使得 ResNet-18 模型更适合处理机器人视觉任务中的高动态场景，提升了模型的实用性和性能。

3.4 噪声调度策略

噪声调度策略在扩散模型中起着核心作用，它决定了如何在不同的迭代阶段向模型引入噪声。正确的噪声调度可以显著影响模型生成输出的质量和稳定性。

3.4.1 噪声调度介绍

噪声调度策略定义了扩散过程中噪声强度的变化方式。这种策略的目标是在模型的训练和推断过程中，逐渐从高噪声数据中恢复出清晰的数据。噪声的逐步减小使得模型能够在每一步更加精确地学习到数据的细节，从而在最终输出时去除噪声，恢复出高质量的结果。

3.4.2 方形余弦调度方法

方形余弦调度是一种在扩散模型中常用的噪声调度技术。这种方法使用一个与方形余弦函数相关的公式来确定每一步中噪声的标准差，提供了一种平滑而有效的方式来减少噪声：

噪声级别 $\sigma(t)$ 通过以下方程进行计算：

$$\sigma(t) = \sigma_{\min} + (\sigma_{\max} - \sigma_{\min}) \times \left(1 - \cos\left(\frac{\pi t}{T}\right)\right)^2$$

其中， t 是当前的时间步， T 是总时间步数， σ_{\min} 和 σ_{\max} 分别是噪声级别的最小值和最大值。

方形余弦调度通过在扩散过程开始和结束时提供更平滑的噪声变化，帮助模型在早期和后期阶段更稳定地学习。这种调度策略的实施确保了扩散过程中噪声的有效管理，使得生成的数据质量更高，且过程更加稳定。通过精确控制噪声的引入，方形余弦调度方法为复杂数据的生成提供了强有力的支持。

4 复现细节

4.1 与已有开源代码对比

为了调整源代码中的参数并成功复现原始研究成果，本文参考了由斯坦福大学实时智能系统实验室发布的https://github.com/real-stanford/diffusion_policy.git。本次作业实现 PushT 任务的机器人操作策略，提供了一套详细的实验框架和参数配置，重现和验证扩散模型在具体机器人任务中的应用效果。

4.2 实验环境搭建

4.2.1 MuJoCo 简介

MuJoCo (Multi-Joint dynamics with Contact) 是一款高效的物理仿真引擎，专为模拟复杂的机械系统和交互动力学而设计。它广泛应用于机器人学、生物力学、计算机图形学以及强化学习等多个领域，其设计重点在于实现高速和精确的仿真。

4.2.2 MuJoCo 的应用意义

1、加速原型开发和测试

- **效率提升：** MuJoCo 通过其高效的物理计算能力，使得研究人员和工程师可以快速迭代和测试各种机器人设计和算法，显著缩短了从概念到实验原型的开发周期。
- **成本节约：** 与物理实验相比，使用 MuJoCo 进行虚拟仿真可以大幅度降低材料和设备的使用成本，特别是在初期探索阶段。

2、提供精确的物理仿真

- **高度精确：** MuJoCo 的高精度物理引擎能够模拟复杂的动力学行为和微妙的接触反应，这对于需要精细操作和高度真实感的应用场景尤为重要，如手术机器人和高级仿生装置的开发。
- **广泛应用：** 准确的仿真结果支持了包括但不限于机器人学、生物力学研究、虚拟现实等领域的深入研究和应用开发，推动了这些领域技术的快速进步。

4.3 数据集介绍：Push-T

Push-T 数据集旨在推进视觉运动策略学习领域的研究，展示了扩散策略在机器人操控任务中的有效性。该数据集由 UR5 机器人捕获，详细记录了平面推动动作，为开发能预测和模拟机器人交互行为的算法提供了宝贵资源。

- **数据组成与格式**数据集存储在与 TensorFlow 兼容的“tfrecord”格式中，包含训练机器人执行复杂操作任务所需的全面数据。这种格式不仅确保了数据处理的效率，还增强了机器学习模型在机器人学中的可扩展性。
- **研究中的应用**该数据集对于完善去噪扩散概率模型（DDPMs）至关重要，便于迭代优化基于多模态感官输入的机器人动作。这些模型的应用在各种机器人基准测试中展示了显著的性能提升，数据集显示平均改进率达到 46.9%。
- **集成到机器人系统中**利用这一数据集可以加速自主机器人的开发，使其能够执行高精度任务。通过在此数据集上进行严格训练，机器人学会了适应复杂环境的动态变化，展现出更优的决策能力。

4.4 实验细节

4.4.1 关键配置和策略

- **策略与模型：**采用 `DiffusionUnetHybridImagePolicy`，结合去噪扩散模型和图像处理技术，优化视觉输入的处理。
- **数据集处理：**使用 `PushTImageDataset` 数据集，专为机器人推动任务设计，含有丰富的动作序列和视觉输入。
- **优化器配置：**利用 `AdamW` 优化器，精调学习率和权重衰减参数，确保模型优化过程的稳定性和高效性。

4.4.2 训练与环境模拟

- **噪声调整策略：**集成 `DDPMScheduler` 进行噪声调整，通过动态控制噪声参数，平衡探索与利用。
- **环境执行器：**配置 `PushTImageRunner` 来模拟机器人在虚拟环境中的操作。
- **性能监控与评估：**定期保存检查点并监控性能，及时调整策略以应对潜在问题。

4.4.3 高级特性与训练优化

- **批处理与工作线程：**设定具体的批次大小和工作线程数，优化数据处理和模型训练的效率。
- **GPU 加速与设备选择：**明确使用 `CUDA` 加速技术，优化模型在 GPU 上的执行性能。
- **学习率调度与优化：**实施余弦退火调度策略，并引入预热步骤。

4.4.4 鲁棒性与恢复能力

- **EMA 策略：**引入指数移动平均（EMA），稳定参数更新，提升模型性能。
- **调试与训练恢复：**支持调试模式和从中断点恢复训练的功能。

5 实验结果分析

1、训练进程分析

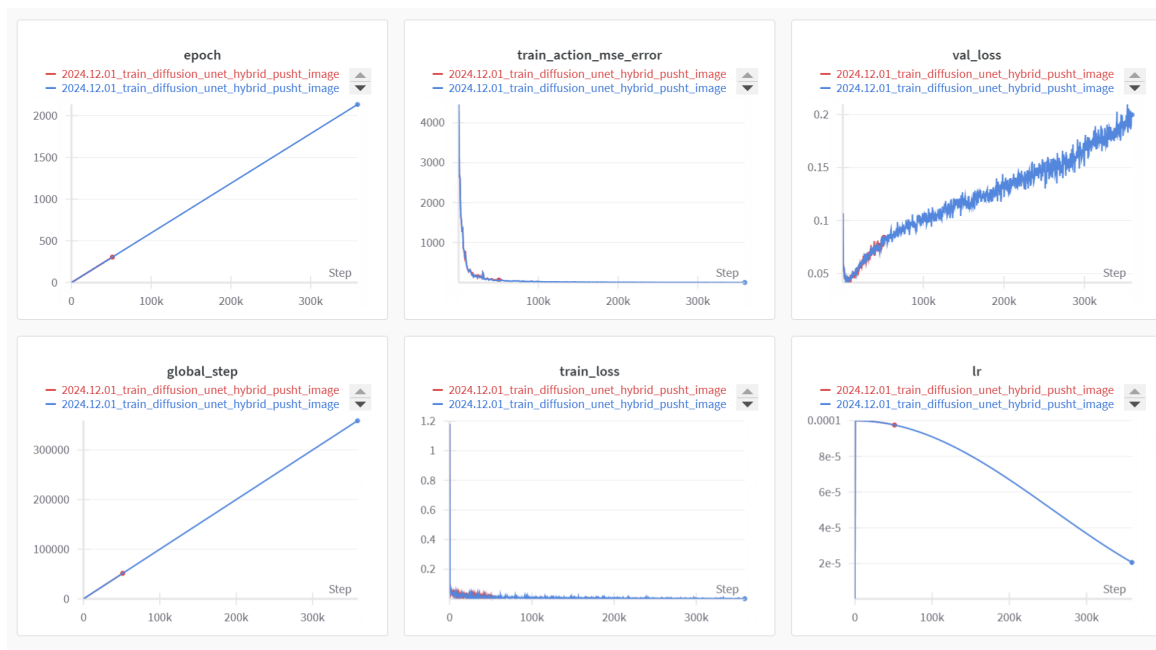


图 2. 实验进程记录

在本次作业中，通过多维度指标系统地监测了一个深度学习模型的训练进程，具体的实验进程如图 2 所示。模型的训练周期（Epoch）和全局步数（Global Step）均匀增长，显示出训练过程的连续性与均匀性。模型在训练早期显示出显著的学习效率，如通过训练动作的均方误差（MSE）迅速下降并快速稳定可见。尽管训练损失（Train Loss）显示出持续的下降趋势，验证损失（Val Loss）的逐步上升却揭示了模型在处理未见数据时的泛化能力受限，这是典型的过拟合现象。此外，实验中采用的指数衰减学习率调整策略，有效地在模型接近最优解时精确调整了参数，以避免过大的参数更新摆动。

本研究结果不仅提供了对模型训练动态的深入理解，还为未来优化算法和提升模型泛化能力提供了宝贵的指导。

2、模型性能评估

在研究中，DiffusionPolicy 方法被用于验证模型在 Push-T 任务中的表现。以下表格详细列出了该策略在不同检测点的成功率：

表 1. 对比原文中的 Diffusion 策略与复现结果在 Push-T 任务的表现

模型	网络架构	准确率
LSTM-GMM (原文)	LSTM-GMM	ph = 0.69 / mh = 0.54
IBC (原文)	IBC	ph = 0.75 / mh = 0.64
BET (原文)	BET	ph = 0.79 / mh = 0.70
DiffusionPolicy-C (原文)	CNN	ph = 0.91 / mh = 0.84
DiffusionPolicy-T (原文)	Transformer	ph = 0.91 / mh = 0.86
DiffusionPolicy-C (复现)	CNN	0.87

考虑到设备上的限制，我选择了 DiffusionPolicy-C 模型，并采用了基于 CNN 的视觉策略。这一选择基于 CNN 在图像处理中的强大能力，特别是在从复杂视觉场景中提取有用特

征方面的优势。从评估结果来看，虽然我的模型复现准确率（0.87）未能达到专业人员（ph）的最佳表现，但是与多人数据集（mh）的结果相近，显示了这一策略在应对真实世界复杂性方面的有效性。

3、模拟结果

下图展示了通过 DiffusionPolicy 方法进行的模拟实验视图，如图所示。这些视图反映了模型在实际操作环境中的表现和交互过程。



图 3. 模拟实验视图

模拟结果验证了模型在执行具体任务时的适应性和准确性，进一步证实了能有效应对现实世界的挑战。

6 未来展望

本研究展示了扩散模型在机器人视觉运动策略学习领域的有效性和潜力，为未来的研究和应用开辟了新的道路。然而，尽管取得了显著成果，目前的方法仍面临一些技术挑战和发展机会。

6.1 技术挑战

- **扩展到更复杂的任务：**当前的扩散模型需要进一步优化以处理更高维度和更复杂的动作序列。未来的研究应致力于提高模型的复杂度和适应性，使其能够在更广泛的机器人操作任务中得到应用。
- **实时性能优化：**虽然模型在离线任务中表现出色，但其在实时应用中的性能还需改进。优化计算效率和降低响应时间将是关键，特别是在需要即时反应的救援或医疗应用中。
- **泛化能力提升：**扩散模型的泛化能力是未来研究的另一个重要方向。通过深入研究和应用先进的算法来增强模型对新环境和未见情况的适应能力。

6.2 发展机会

- **工业自动化：**随着工业 4.0 的推进，扩散模型可用于开发更智能的自动化机器人，特别是在高精度的制造和组装线中。
- **服务机器人领域：**在服务行业，如餐饮和医疗服务中，扩散模型能够提供更为精细和个性化的服务，提升客户体验和服务效率。
- **机器人教育和互动应用：**教育机器人可以利用扩散模型来提供更加互动和引人入胜的学习体验，尤其是在语言学习和技能培训上。

参考文献

- [1] Brenna D Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, 2009.
- [2] Lili Chen, Kevin Lu, Aravind Rajeswaran, Harrison Lee, Aditya Grover, and Igor Mor-datch. Decision transformer: Reinforcement learning via sequence modeling. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.
- [3] Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. In *Proceedings of Robotics: Science and Systems (RSS)*, 2023.
- [4] Yihe Dong, Jean-Baptiste Cordonnier, and Andreas Loukas. Attention is not all you need: Pure attention loses rank doubly exponentially with depth. *IEEE Transactions on Neural Networks*, 139, 2021.
- [5] Peter Florence, Lucas Manuelli, and Russ Tedrake. Self-supervised correspondence in visuomotor policy learning. *Journal of Robotics Research*, 5:492–499, 2020.
- [6] Jonathan Ho and Stefano Ermon. Generative adversarial imitation learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 4565–4573, 2016.
- [7] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *arXiv preprint arXiv:2006.11239*, 2020.
- [8] Tianyi Jin and Jingjing Zhang. Generating adaptive robotic behaviours via enhanced diffusion policy. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2024.
- [9] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. In *International Conference on Learning Representations (ICLR)*, 2014.
- [10] Xiaoming Lin, Wen Xu, and Yuxuan Mao. Vision-based tactile image generation via contact condition-guided diffusion model. *arXiv preprint arXiv:2412.01639*, 2024.

- [11] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2013.
- [12] Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *International Conference on Machine Learning (ICML)*, 2021.
- [13] Ethan Perez, Florian Strub, Harm de Vries, Vincent Dumoulin, and Aaron Courville. Film: Visual reasoning with a general conditioning layer. *IEEE Transactions on Neural Networks and Learning Systems*, 32, 2018.
- [14] Michael Reuss, John Pari, and Pradeep Agrawal. Efficient diffusion transformer policies with mixture of expert denoisers for multitask learning. In *Workshop on Mastering Robot Learning, OpenReview*, 2024.
- [15] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. *arXiv preprint arXiv:2112.10752*, 2022.
- [16] Chitwan Saharia, Jonathan Ho, William Chan, Tim Salimans, David J Fleet, and Mohammad Norouzi. Palette: Image-to-image diffusion models. *arXiv preprint arXiv:2111.05826*, 2022.
- [17] John Schulman, Sergey Levine, Philipp Moritz, Michael I Jordan, and Pieter Abbeel. Trust region policy optimization. In *International Conference on Machine Learning (ICML)*, pages 1889–1897, 2015.
- [18] Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.
- [19] Zhiming Wang, Zhi Li, Animesh Mandlekar, and Zhengjia Xu. One-step diffusion policy: Fast visuomotor policies via diffusion distillation. *arXiv preprint arXiv:2410.21257*, 2024.
- [20] Haochen Zhu, Tianhao Zhao, and Ludovic Righetti. Should we learn contact-rich manipulation policies from sampling-based planners? *arXiv preprint arXiv:2412.09743*, 2024.