



北京大学

# 本科生毕业论文

题目： 利用多来源机器人数据进行表征学习

---

**Exploiting Effective Representations**

---

**via Cooperative Learning of**

---

**Multi-Sensory Robotics Data**

---

姓 名： 倪超

---

学 号： 1500011042

---

院 系： 工学院

---

专 业： 理论与应用力学

---

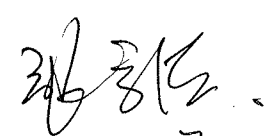

指导教师： 张崇洁 助理教授 唐少强 教授

---

二〇一九年六月

# 北京大学 工学院

## 本科生毕业论文评语及成绩

学号	1500011042	姓名	倪超	专业	理论与应用力学
导师姓名	张崇洁	导师职称	助理教授	导师单位	清华大学交叉信息研究院
导师姓名	唐少强	导师职称	教授	导师单位	工学院
论文题目	中文	合作利用多来源机器人数据进行表征学习			
	英文	Exploiting Effective Representations via Cooperative Learning of Multi-Sensory Robotic Data			
指导教师评语	<p style="font-size: 1.2em;">论文选题有意义, 也很有挑战, 对前表征学习提出了一些新的想法, 文献材料详实, 实验设计合理, 并取得很意义的结果. 证明该学生具有很强的理论基础和科研潜力.</p> <div style="text-align: right; margin-top: 20px;">               指导教师签字:              2019年5月24日         </div>				
论文成绩	93	答辩小组负责人签字			

# 合作利用多来源机器人数据进行表征学习

## 摘要

从高维数据中挖掘有效的信息是深度学习中一个核心问题,在强化学习中采集到的信息都是没有标记过的,因此在过去的强化学习算法中,一个普遍的方法是利用预测将来数据进行自监督学习。机器人在执行任务时,往往会涉及到多种类型的数据,比如机器人控制状态,相机观测结果等等。这些数据能够从多方面帮助我们认识机器人的运动状态。在这篇文章中,我们提出了一个新的合作化学习框架以尽可能从多种机器人数据中提取有效信息。我们的模型将自监督预测学习和同步化学习结合在一起,综合利用了各种数据流各自的动力学信息,以及不同数据之间的时序上的一致性。我们在流行的机器人控制平台上应用我们的模型,结果展示我们的模型能够有效的提取出数据中的高级特征,这些特征能够应用于机器人强化学习问题中,且模型具有鲁棒性和泛化性。

**关键词:** 无监督学习, 同步化学习, 预测学习

# Exploiting Effective Representation via Cooperative Learning of Multi-Sensory Robotics Data

Chao Ni (Theoretical and Applied Mechanics)

Supervised by Chongjie Zhang, Shaoqiang Tang

## Abstract

Extracting effective representations from high-dimensional data has been one of the core problems for deep learning. Since the collected data by a reinforcement learning agent are unlabeled, previous reinforcement learning methods learn task-dependent representations in the policy learning process or task-agnostic representations by predicting the future in the pre-training. In robotics tasks, multi-sensory data (e.g., robotic control state and camera observation) are usually accessible and provide abundant multi-source information. In this paper, we present a novel cooperative learning framework to extract effective representations from multi-sensory data. Our framework integrates self-supervised predictive dynamics learning and contrastive synchronization learning, which respectively leverages the dynamics information of each data source and the natural temporal synchronization of multi-source data. We evaluate our method on popular robotic control tasks, and results show that our model learns effective generic representations for down-stream reinforcement learning tasks and enables efficient transfer among multi-sensor data.

**KEYWORDS:** Unsupervised learning, Synchronization, Prediction learning

# 目 录

第一章 引言.....	1
第二章 自监督预测和同步化.....	2
2.1 多来源数据自监督预测和同步化.....	2
2.2 损失函数.....	2
2.3 负样本选取.....	4
2.4 网络模型.....	5
2.4.1 SCN-SINGLE.....	6
2.4.2 SCN-DOUBLE.....	6
2.4.3 SCN-SKIP.....	6
2.4.4 SCN-LSTM.....	6
第三章 实验设定.....	7
第四章 实验结果和讨论.....	8
4.1 预测和同步结果展示.....	8
4.2 预测对于同步化的影响.....	9
4.3 同步化对预测的作用.....	11
第五章 讨论与结论.....	12
参考文献.....	13
致谢.....	15

# 第一章 引言

从高维的标记数据中进行表征学习以提取数据的高级特征是人工智能的一个重大成功。深度神经网络的出现为表征学习提供了一个十分宽广的平台，依靠神经网络，处理海量级的数据并从中学习特征成为可能。He, Kaiming, et al [2016] 提出了 Resnet、Simonyan, Karen, and Andrew Zisserman [2014] 提出 VGG、Jürgen Schmidhuber [1997] 提出 LSTM, 这些网络为深度学习平台提供了有力的工具，并且推动了各种最新机器学习工具的发展与应用。但是，依然有许多问题有待解决，比如在机器人背景中，数据相对有限，如何有限利用数据是一个主要的问题。Deisenroth, et al [2011] 提出了基于模型的算法，能够解决在相对小的数据量下的机器人控制问题。同时，有效的深度学习环境还要求网络具有良好的鲁棒性和泛化性。

改善表征学习的效果要求我们能够学习到一些与具体对象无关的特征。比如，当我们在学习图像分类的时候，通过在 Mubarak Sha, et all [2012] 采集的 UCF101 上学习得到的分类器应该能够在 Shahroudy, Amir, et al [2016] 收集的 NTU 数据集上同样起效果，同时，通过分类这一有监督模式下学习到的分类器往往会忽视一些别的信息，通过分类问题预训练的网络在图像命名方面的表现必然会下降。也就是说，有监督的学习是有偏误的，并且这个偏误是和具体任务背景相关的。因此，无监督学习是提高表征学习的泛化性的重要一步。

一个常见的无监督学习的策略是预测将来的信息。Peter, et al [1955] 在信号处理和数据压缩中利用预测进行无监督学习；近年来无监督学习中，Tomas, et al. [2013] 已经成功利用了这些想法，并且利用预测临近单词的方式成功实现了词汇中的表征学习。此外，利用不同数据之间相关性进行自监督学习来学习特征也是目前流行的一种方法，Aytar, et al. [2016], Arandjelovic, et al. [2017] 利用了视频和音频之间的关联性应用了自监督学习机制。Korbar, et al. [2018] 利用视频和音频之间的同步化损失提取出高级特征并应用到了动作识别任务中。

在这篇文章中，我们提出了一种自监督学习框架，能够将动态预测学习和同步化学习结合起来，综合利用不同来源数据，提取出任务不敏感的特征提取方法。这种框架网络能够实现同步化，而在同步化中学习的知识能够用于强化学习任务。

## 第二章 自监督预测和同步化

这一部分，我们会给出同步化损失的具体含义，并给出模型训练数据的具体收集方式。并且提出用来结合自监督预测和同步化的网络实现形式。

### 2.1 多来源数据自监督预测和同步化

我们假设给定有训练集  $D = \{(I^{i}, v^{i}, y^{i}, a^{i}, \tilde{I}^{i}, \tilde{v}^{i})\}$ ，训练集由  $N$  对图像-向量对组成。在这里  $I^{i}, v^{i}$  分别表征图像和向量的一段序列。而  $\tilde{I}^{i}, \tilde{v}^{i}$  是在图像向量对之后时序上的下一组数据。 $y^{i} \in \{0,1\}$  是用来表征这一对图像-向量组合是否是截取于同一段动作。也就是说，如果他们取样于同一段动作视频的相同时间段，那么我们说这个样本是同步的， $y^{i} = 1$ ；相反，如果他们来自于不同动作，或者是同一段动作的不同时间点，那么就认为他们没有同步， $y^{i} = 0$ 。我们的任务之一，便是需要学习一个函数，或者说编码器  $g(x)$ ，使得  $g(I^{i}, a^{i}) = y^{i}$  在尽可能多的样本上满足。而由于图像和向量本身数据的特点，我们将对图像和向量分别处理，因此实际的同步化处理判断将在编码之后处理。若记两个编码器分别为  $f_I(x), f_v(v)$ ，则针对同步化的问题可以等价于：

$$\max_{f_I, f_v} \sum_{i \in S} I\{g(f_I(I^{i}), f_v(v^{i})) = 1\} \quad (1)$$

其中， $I\{\cdot\}$  是一个逻辑函数。 $S$  是所有同步化的样本集合。在具体计算时，我们会引入等价的损失函数作为优化对象。

### 2.2 损失函数

一个常见的选择是利用交叉熵来作为损失函数，因为这个训练过程本质上是一个二分类的问题，但是受到 Korbar, et al. [2018] 影响，我们发现采取对比损失作为损失函数具有更加好的收敛性能和鲁棒性能。在我们的设定里，我们选取参数使得同步的图像-向量对能够尽可能接近，而不同步的图像-向量对能够尽可能远离。因此我们选取的损失函数具有如下的形式：

$$\begin{aligned} Loss_{syn} = & \frac{1}{N} \sum_{n=1}^N (y^{n}) \|f_I(I^{n}) - f_v(v^{n})\|_2^2 \\ & + (1 - y^{n}) \max(\gamma - \|f_I(I^{n}) - f_v(v^{n})\|_2, 0)^2 \end{aligned} \quad (2)$$

其中,  $\gamma$  是一个边际超参数, 一般会取一个小于 1 的正实数。而对于在新数据集上的预测, 可以定义  $g(f_I(I^{[n]}), f_v(v^{[n]})) = I\{\|f_I(I^{[n]}) - f_v(v^{[n]})\|_2 < \tau\}$ 。若隐空间上两者距离小于阈值  $\tau$ , 那么可以认为图像-向量对是同步的。

同时, 在实验中, 我们也认为, 相邻的图像序列之间的距离应该比远程的图像序列更近, 向量序列也是如此, 因此我们在 SCN-CTR 中考虑了图像-图像对和向量-向量对之间的对比损失。这在本质上其实是优化互信息的一种手段 Oriol Vinyals, et al [2018]。引入内部数据集  $F = \{(I^{[i]}, I^{[j]}, y^{\{i-j|=1\}}), (v^{[i]}, v^{[j]}, y^{\{i-j|=1\}})\}$ 。如果图像对在时间上是邻近的, 那么认为其为正例, 否则为反例。考虑时间顺序上的一致性之后, 我们可以引入同类之间的对比损失:

$$\begin{aligned} Loss_{syn1} = & \frac{1}{N} \sum_{i,j} (y^{\{\delta\}}) \|f_I(I^{[i]}) - f_I(I^{[j]})\|_2^2 \\ & + (1 - y^{\{\delta\}}) \max(\gamma - \|f_I(I^{[i]}) - f_I(I^{[j]})\|_2, 0)^2 \\ & + \frac{1}{N} \sum_{i,j} (y^{\{\delta\}}) \|f_v(v^{[i]}) - f_v(v^{[j]})\|_2^2 \\ & + (1 - y^{\{\delta\}}) \max(\gamma - \|f_v(v^{[i]}) - f_v(v^{[j]})\|_2, 0)^2 \end{aligned} \quad (4)$$

其中,  $\delta = I\{|i - j| = 1\}$  为逻辑函数。

与 Korbar, et al. [2018] 不同的是, 我们认为在同步化过程中引入对下一帧图像的预测能够帮助同步化是可控的。在编码器将图像和向量映射到隐空间之后, 我们需要保证隐空间中依然含有足够多的信息量, 这些信息量在给予动作之前能够重构成原有图像或向量, 并且能够在施加动作之后预测到下一帧的图像。因此我们引入了预测损失如下:

$$Loss_{pred} = \frac{1}{N} \sum_{n=1}^{N-1} (\|I^{[n+1]} - h_I(f_I(I^{[n]}), a^{[n]})\|_2^2 + \|v^{[n+1]} - h_v(f_v(v^{[n]}), a^{[n]})\|_2^2) \quad (5)$$

其中,  $h(l, a)$  为解码器, 将隐空间作用动作之后重新映射到图像或者向量。因此, 我们所采取的整体损失函数为同步化损失和预测损失的加权之和。为了证明引入预测损失函数对于训练同步性解码器的作用, 我们将在实验部分比较只训练  $Loss_{syn}$  和同时训练  $Loss_{syn} + Loss_{pred}$  的差异。



## 2.3 负样本选取

在训练过程中，我们选取等量的正样例和负样例。其中正样本容易选取，Mojuco 环境生成的图像和向量本身便是相对应的，我们只需要取相同长度的序列即可。我们从 Mojuco 生成的视频序列中随机选取长度固定的片段作为正样本。而对于负样本，我们又做了细分：困难和超困难。其中，困难样本是取自同一段视频但是没有重叠的片段，而超困难样本则是取自同一段视频但是同时有部分重叠的情况。图 1 给出了我们选取样本的方式举例。

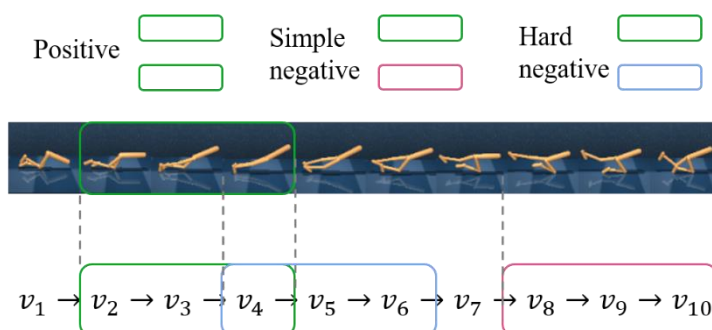


图 1: 样本选取方式举例。正样本是在不同数据流上有相同时间位置的数据对，简单负样本是在不同位置而且无重叠的数据对，难负样本是部分重叠的数据对。

而在考虑图像和向量本身之间的临近关系之时，我们给出正样本和负样本如图 2 所示：



图 2: 同类数据流临近样本和远离样本举例。临近样本是在时间上严格临近的数据对，而远离样本是任何非严格临近的数据对。

我们在选取的时候，首先不考虑最优化互信息的情况，我们选取正样本和负样本各一半，并且在负样本中选取难样本和超难样本也各一半。显然，重叠的样本数目越多，负样本于正样本就越难以区分开，我们在后文也会介绍不同比例的难样本和超难样本对网络学习能力的影响。

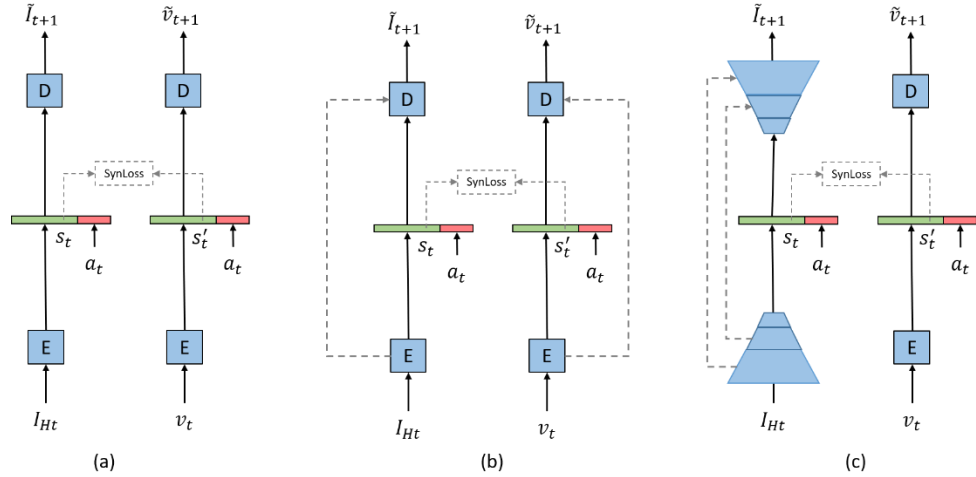


图 4: SINGLE, DOUBLE 和 SKIP 模型示意图。(a) 为 SINGLE 情形，数据经过编码器之后只有一条通道，编码之后经过同步化并且施加动作之后，经过解码器得到下一个数据的预测。(b) 是 DOUBLE 的情形，也就是说数据经过编码器之后又两条通道，其中一条是与同步化和动作无关的信息；(c) 是 SKIP 的情形，本质上也是一种 DOUBLE 类型的网络，它在编码之时会将一些低级特征直接传送到解码器之中。

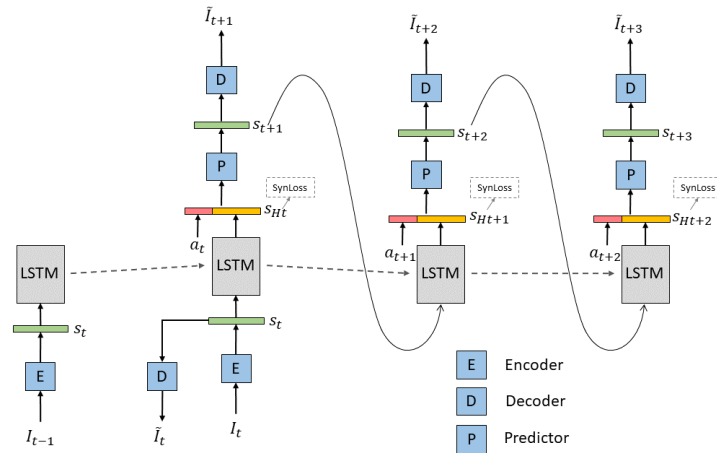


图 3: LSTM 模型示意图。利用历史信息预测将来的若干时间点的信息。数据在经过编码器之后，输入到 LSTM 单元，结合单元内的历史信息之后，经过同步化和环境动作输送到解码器得到下一数据的预测，同时这部分也作为下一个 LSTM 单元的输入来预测往后更多的数据信息。

## 2.4 网络模型

针对预测问题和同步问题，我们的基本思想如下，第一步是同步化问题，首先对于输入的图像和向量，分别进入编码器，对于图像和向量，我们采取了不同的编码器。期中图像的编码器是由五层卷积神经网络和一层全连接网络组织，这一网络结构参考了文献 Danijar, et al. [2018] 而向量的编码器是一个三层全连接和一层激活函数组成的神经网络。图像和向量经过各自的编码器之后映射到一个共同的隐空间，通过计算在隐空间两者的距

离可以计算同步化损失。第二步是预测问题。由于下一帧的图像是基于当前帧以及当前时间的动作的，在网络中我们需要在合适的地方加入动作。我们选择在隐空间上对图像和向量施加动作函数。我们认为，图像和向量共同的隐空间是能够表征图像和向量一致性的空间，也就是能够过滤到一些无关背景信息，一些不随着外界环境改变的信息，因此隐空间是一个精准的机器人的运动空间。在隐空间给予动作之后，便得到了施加下一帧图像或者像一个向量的隐空间，经过解码器之后，便能够得到下一帧的预测值。

在具体实现过程中，我们又考虑各种变化，以下将分别介绍。

#### 2.4.1 SCN-SINGLE

这便是前文所叙述的基本模型，图像经过编码器之后映射到隐空间，在隐空间施加动作之后再经过解码器映射到原图像空间，得到对下一帧图像的预测。其网络结构如图 3(a)所示。其中， $I_{Ht}$  表示利用长度为  $H$  的图像预测下一帧。在我们的情形中， $H$  取为 4，也就是说在每一次预测中，利用了前面四帧图像的信息。

#### 2.4.2 SCN-DOUBLE

SCN-DOUBLE 模型是对于 SINGLE 的改进。我们注意到，图像和向量即使都是对于一个机器人运动的描述，但是其所包含的信息还是不同的。对于机器人而言，其向量信息包含的为关节的三维坐标，特征点离标志点的距离等等。而图像信息为一个平面信息，包含更多无关的背景像素。与此同时，向量比图像多了三维信息。因此我们认为，经过同步化之后，图像在捕捉到了运动信息的同时，也失去了许多背景信息等与机器人动作无关的信息。因此 double 模型考虑了这一部分与运动无关的信息，这部分信息经过编码器编码之后不受动作影响，而直接经过解码器解码预测下一帧的图像。具体模型见图 3(b)。

#### 2.4.3 SCN-SKIP

SCN-SKIP 与 SCN-DOUBLE 类似，不仅仅利用了隐空间上的信息，而与前一种模型的不同在于，前者利用了神经网络中 highway 的思路，也就是说，在编码过程中，将一些中间层的结果直接连接到解码器中。在解码过程中，从第一步的图像到最后的隐空间，其低维度的信息逐渐被剥离，而高级的隐式的信息逐级保留。诸如图像的背景，轮廓分布等等低级别信息对于预测下一帧图像却可能有更为显著的效果。这个模型见图 3(c)。

#### 2.4.4 SCN-LSTM

在之前的模型中，我们都采用了长度为  $H$  的图片和向量来预测下一帧，但采用的都是基本的 CNN 的网络。针对历史信息，有一种更好的模型是 LSTM(long short-term memory) Jürgen Schmidhuber [1997]。这是一种常见于自然语言处理中的方法，近年来也常常应用

于视频的动作识别等领域。在 LSTM 网络中，每一个 LSTM 单元中都会存有历史信息，并且会在不断输入中更新历史信息。在我们构造的 LSTM 模型中，我们利用历史的四帧图像作为输入来预测之后的四帧图像。对每一帧图像，先经过编码器输入 LSTM 单元之后，LSTM 单元会更新历史信息，而当前帧在隐空间接受动作之后经过预测得到下一帧的隐空间映射，该映射便被作为对之后一帧图像预测的输入。具体网络如图 4 所示。

## 第三章 实验设定

在训练 SCN 网络的时候，我们选取了 500 个视频，每一个视频的长度不超过 1000 帧，对图像和向量我们均做归一化处理。所有图像被缩小到  $64*64*3$ 。我们的网络在 Geforce GTX 1080 Ti 上运行。经过测试，我们发现运行 400 轮后所有网络都能达到收敛。对于参数选择，我们选取超参数  $\gamma = 0.99$ ，隐空间维度对于非 LSMT 情况都选为 32 维，而对于 LSTM 情况，实际在隐空间上的维数为 256 维。在运算过程中，我们给予同步化损失，图像预测损失和向量预测损失不同的权重。

# 第四章 实验结果和讨论

## 4.1 预测和同步结果展示

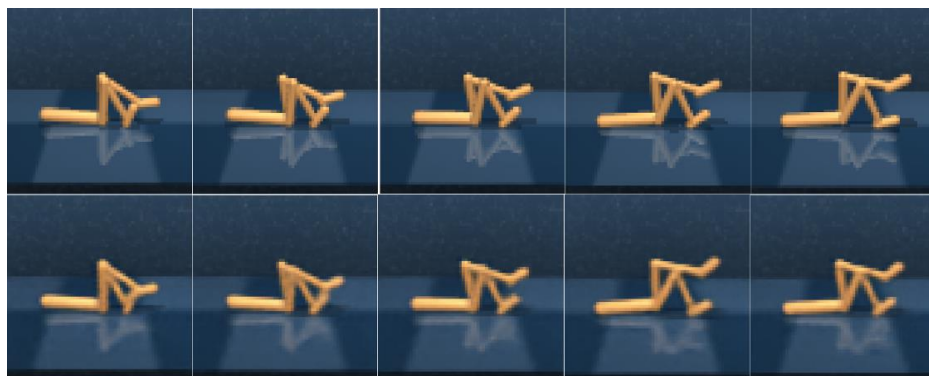


图 5: 机器人预测结果展示。在图中，上面一排是真实的机器人动作，而下面一排是预测得到的结果。可以看到，网络已经成功预测到了细微的动作变化，比如脚尖的动作方向等等。

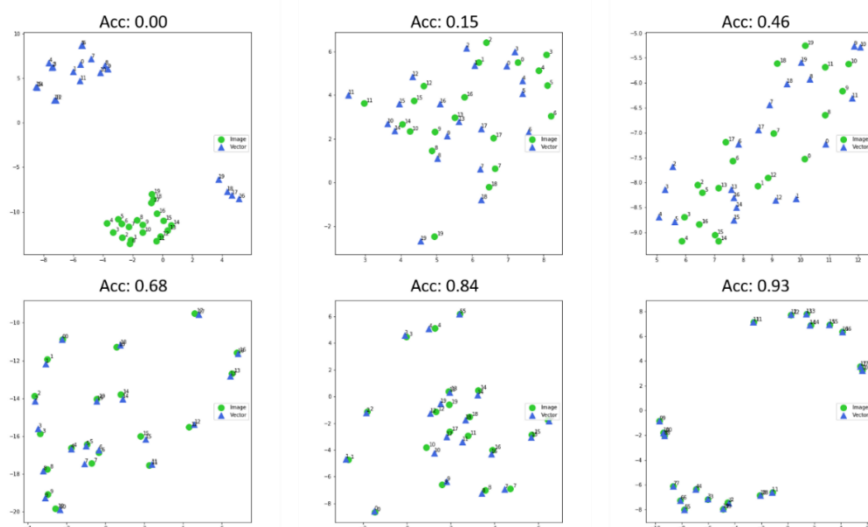


图 6: 不同同步化的可视化结果以及相应的同步化准确度比较。上图为隐空间向量中在二维平面的可视化展示。在隐空间越接近的向量，在经过 TSNE 可视化处理后的二维空间也越接近。在上图中，蓝色表示向量，红色表示图片。我们选取了连续二十个点来进行可视化处理。第一幅图为尚未经过训练时的可视化效果，此时图片信息和向量信息在随机解码下没有同步，相反，由于各自用了同一种解码器，图片和向量内部产生了集聚。随着训练进行，图片和向量逐渐同步化，同步准确率也逐渐提高。

我们考虑了图像预测损失、向量预测损失和同步化损失，并设定学习率为 0.0001。经过四百轮学习之后，我们在各种不同的任务上训练网络，得到各种方法的结果如表 1 所示。从表中我们可以看到，SCN-DOUBLE 模型能够实现最好的同步化效果，其次是利用

LSTM 模块也能够实现更好的效果。注意到 LSTM 的图像预测损失和向量预测损失是预测将来四帧图像的预测损失之和，而越远离当前帧的图像预测由于不知道之后的动作分布，其预测是相对随机的，因此其预测损失也偏大。接下来我们给出预测的实际效果图，如图 5 所示。其中，上面一排是真实的图像分布序列，而下面一排是经过预测解码之后生成的图像分布。可以看到，训练到四百轮之后，已经能够预测到机器人的微小动作变化。对于同步化损失，我们将高维的隐空间向量可视化到二维平面，在二维平面的距离大小可以表征高维隐空间中的远近。我们给出了不同同步化损失下的二维可视化示意图如图 6 所示。在图中，我们随机选取了在同一个视频中的连续二十个点的图像和向量数据，绿色点表示图像，蓝色三角形表示向量。其中第一幅图是在尚未开始训练的时候，图像和向量在隐空间上的分布。我们可以看到，绿色和蓝色点之间相隔较远，且形成了明显的分区。这说明在随机初始化的编码器作用下，不能够得到同步化的结果。在经历四百轮训练之后，我们可以看到，蓝色点和绿色点之间对应于同一个视频序列点的距离在缩小，而不对应的点之间开始相互分离。图中最后一幅图是同步化损失在 0.04 左右的情况，可以看到此时图像和向量已经充分对齐。但是从图 6 也可以看到，不同时间点之间的分布是随机的，也就是说，我们虽然捕捉成功将图像和向量映射到了同一个隐空间，但是我们的映射是孤立的，也就是映射中不包含帧与帧之间的时间关系。为了捕捉帧与帧之间的关系，我们引入了同种变量之间的对比损失函数。如方程 4 所示。作为比较，我们给出了是否引入同方向变量损失函数的隐空间映射比较，如图 7 所示。图中左边是 SCN-DOUBLE 模型在进行三百轮训练之后得到的映射图像，而右边是引入同变量同步损失函数之后的映射图像。可以看到，引入同变量之间同步损失函数之后，在原空间具有一定顺序的图像和向量对，在隐函数空间依然具有一定的顺序。

## 4.2 预测对于同步化的影响

我们认为，尽管 Korbar, et al. [2018] 只引入了同步化损失对多来源机器人数据进行了同步化，并取得了很好的效果。但是，这种在隐空间同步化的方式没有受到约束。也就是说，编码器可能会出现这样一种情况，它找到了多来源数据的一个十分微小的共同点，而这个共同点也许仅仅是某种形状、颜色等等并不是十分高级的特征，或者仅仅是部分高级特征，而这些特征已经足够用来做同步化，但是却没有包含着所有的有意义的信息。因此，我们认为，对隐空间中的元素做一个约束是必要的。我们要求，映射到隐空间之后，数据依然需要包含足够的信息，我们引入对预测的要求，也就是说，数据在施加动作之后能够和下一个时间点上的数据一致，能够预测下一个时间点的信息。我们进行了只含有同步化损失和兼有同步化损失和预测损失的比较，结果发现，如果仅仅考虑同步化损

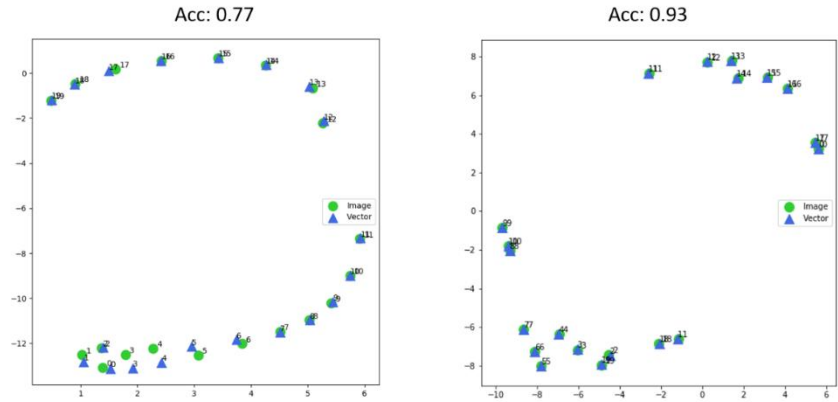


图 7: 引入同种数据的对比损失函数之后的 TSNE 可视化结果比较图。图右是没有引入的情况，可以看到点的分布是没有规则的，而引入之后，本具有时间顺序的点在隐空间也具有一定的时间顺序。如左图所示。

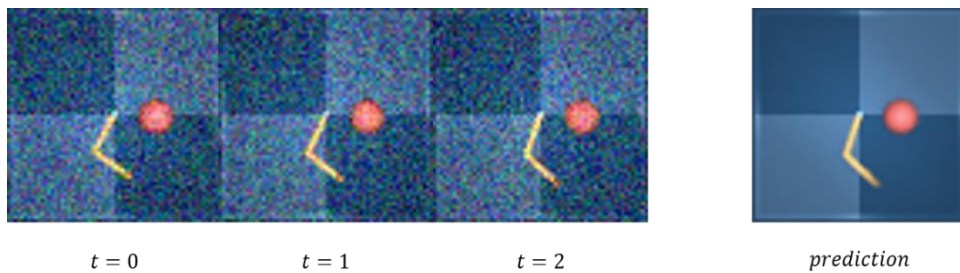


图 6: 预测对图像噪声的免疫性。上图中左边三幅为连续时间上的三幅图，而右边是预测得到的下一帧图像。其中左图在输入时人为引入了噪声，但是结果得到预测能够免疫噪声。

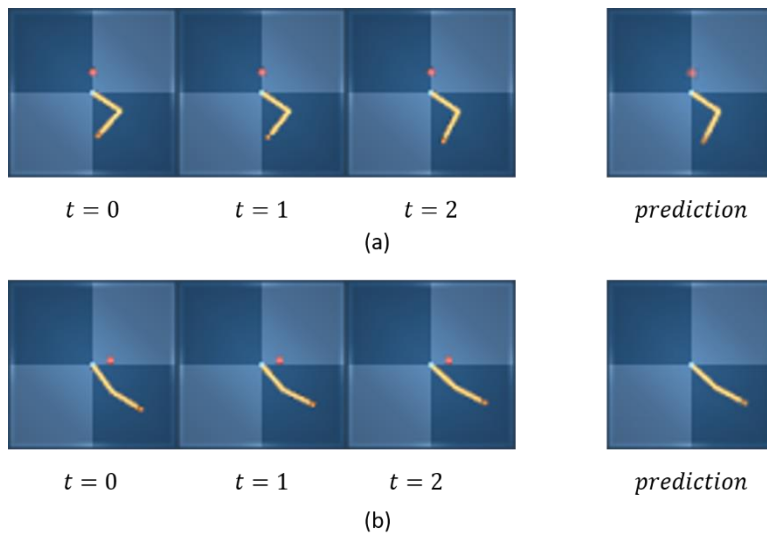


图 5: 同步化对预测的作用。图中(a)中为引入同步化之后，通过历史信息预测到下一数据的展示，可以看到，预测得到的信息包含小球信息。而(b)中由于没有引入同步化，且小球本身作为背景很小，因此无法预测到小球的存在。

失，那么在训练过程中十分容易出现过拟合现象。也就是说，网络学到了十分有限的知

识，而即使这些知识已经足够用来保证信息之间的同步，这些信息不能够充分来表征机器人的运动。此外，由于真实生活中数据会带有许多噪声，因此我们测试了在携带有噪声下的数据是否能够在网络中实现预测和同步化，得到结果如图 8 所示：

### 4.3 同步化对预测的作用

在不同来源机器人数据之间，其包含内容不会完全相同。比如图像信息只能包含某一个平面信息，而失去了立体的三维信息，同时图像还包括大量的背景信息，而背景信息是向量所不具有的。通过不同来源数据之间的同步化，我们相信在共同的隐空间上，不同数据会相互学习到各自不含有的信息。为了验证这个想法，我们选取了 **Reacher** 这一任务，并调整了困难模式，也就是红球形状很小的情况。该任务中机械臂需要碰撞红球来获得奖励。如果不考虑同步化，仅仅依靠图像本身去预测的话，由于红球本身很小，并且在股东的轨迹中是犹如背景一样保持不动的，我们发现网络很难预测到红点的存在。而当我们考虑同步化损失之后，由于在向量中包含球到中心点的距离，在同步化过程中，隐空间中的高级特征同时会兼有图像和向量的信息，通过同步化，球的位置信息会被学习到，并帮助到对下一节点数据的预测。比较图如图 9 所示。

表 1：模型在不同任务上的同步化结果。在所有模型中，LSTM 模型能够取得最好的效果，但是同时需要更多的训练时间。我们在三个场景进行训练：walker stand, easy reacher, hard reacher。并且在其中一个场景加上了噪音，在另一个场景人为减去了部分数据。

<i>Method</i>	<i>Task</i>	<i>Epochs</i>	<i>Accuracy(%)</i>
SCN-SINGLE	Walker: stand	1-300	74.7
	Reacher: easy (noise)	1-300	46.8
	Reacher :hard	1--300	95.4
SCN-DOUBLE	walker: stand	1-300	84.5
	Reacher: hard	1-300	96.9
SCN-SKIP	Walker: stand	1-300	45.3
	Reacher: easy (cut)	1-300	80.7
	Reacher: hard	1-300	96.9
SCN-CTR	Walker: stand	1-300	24.9
SCN-LSTMSINGLE	Walker: stand	1-400	92.9
SCN-LSTMSKIP	Walker: stand	1-400	98.3



## 第五章 讨论与结论

在这篇文章中，我们提出了一个新的合作化学习框架以尽可能从多种机器人数据中提取有效信息。我们的模型将自监督预测学习和同步化学习结合在一起，综合利用了各种数据流各自的动力学信息，以及不同数据之间的时序上的一致性。通过在强化学习机器人平台上测试，我们的模型得到了验证。但是还有有待解决的地方。我们尚未将我们的编码器应用于强化学习任务，我们需要证明同步化模型在结合多数据信息之后，能够显著改善强化学习任务，并广泛用于各种机器人背景实验环境。

## 参考文献

- He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
- Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780.
- Deisenroth, Marc, and Carl E. Rasmussen. "PILCO: A model-based and data-efficient approach to policy search." *Proceedings of the 28th International Conference on machine learning (ICML-11)*. 2011.
- Soomro, Khurram, Amir Roshan Zamir, and Mubarak Shah. "UCF101: A dataset of 101 human actions classes from videos in the wild." *arXiv preprint arXiv:1212.0402* (2012).
- Shahroudy, Amir, et al. "NTU RGB+ D: A large scale dataset for 3D human activity analysis." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- Elias, Peter. "Predictive coding--I." *IRE Transactions on Information Theory* 1.1 (1955): 16-24.
- Mikolov, Tomas, et al. "Efficient estimation of word representations in vector space." *arXiv preprint arXiv:1301.3781* (2013).
- Aytar, Yusuf, Carl Vondrick, and Antonio Torralba. "Soundnet: Learning sound representations from unlabeled video." *Advances in neural information processing systems*. 2016.
- Arandjelovic, Relja, and Andrew Zisserman. "Look, listen and learn." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.

Korbar, Bruno, Du Tran, and Lorenzo Torresani. "Cooperative learning of audio and video models from self-supervised synchronization." *Advances in Neural Information Processing Systems*. 2018.

Oord, Aaron van den, Yazhe Li, and Oriol Vinyals. "Representation learning with contrastive predictive coding." *arXiv preprint arXiv:1807.03748* (2018)

Hafner, Danijar, et al. "Learning latent dynamics for planning from pixels." *arXiv preprint arXiv:1811.04551* (2018).

# 致谢

首先，我要感谢我的毕业论文导师张崇洁，他带我进入了强化学习这一领域，通过在他组里的学习，我逐渐明确了自己将来要从事的方向。其次，我也要感谢我大三时在北大的导师王启宁和在约翰霍普金斯大学的导师 Gregory Chirikjian，他们逐渐引领我发现自己兴趣，并且在各方面给予我充分的建议和指导。第三，我还要感谢在毕业论文完成过程中和我合作的朱广翔师兄，他给予了我第一手的做科研的经验，让我从细节处逐渐掌握做科研的能力。