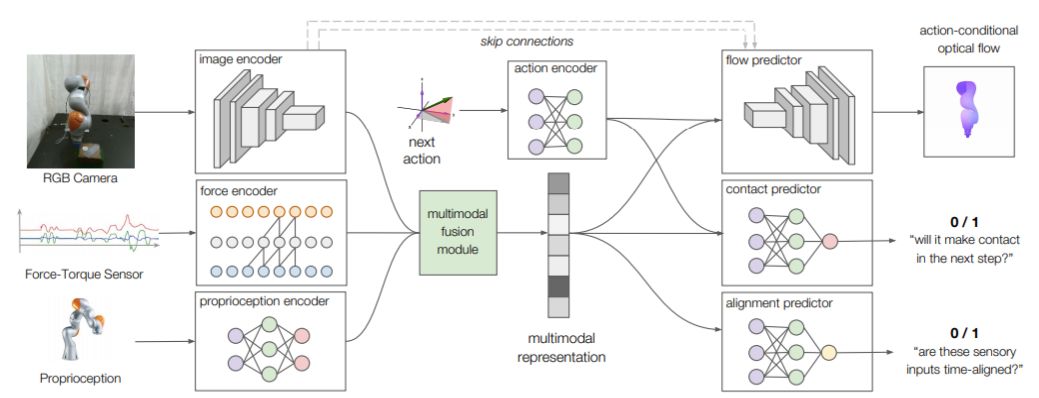
**Best paper in ICRA2019**

**论文 ：Making Sense of Vision and Touch: Self-Supervised Learning of Multimodal Representations for Contact-Rich Tasks**

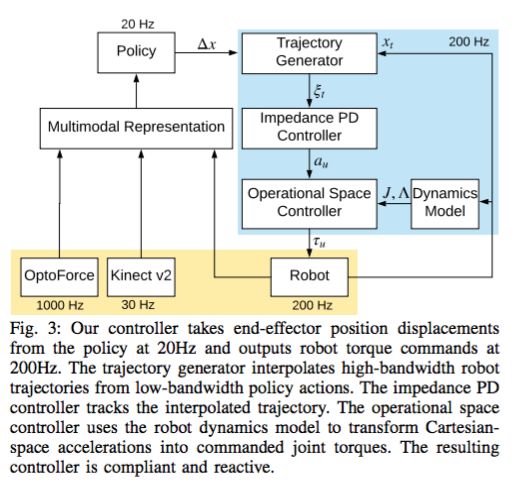
该研究使用自监督学习感知输入的紧凑、多模态表征，然后使用这些特征提升样本效率。研究者在植入任务上评估了该方法，结果表明该方法对于外部扰动具备稳健性，同时可以泛化至不同场合。此外，研究者展示了在模拟环境中和真实机器人上的结果。

该研究提出的多模态表征学习模型架构如下图所示：

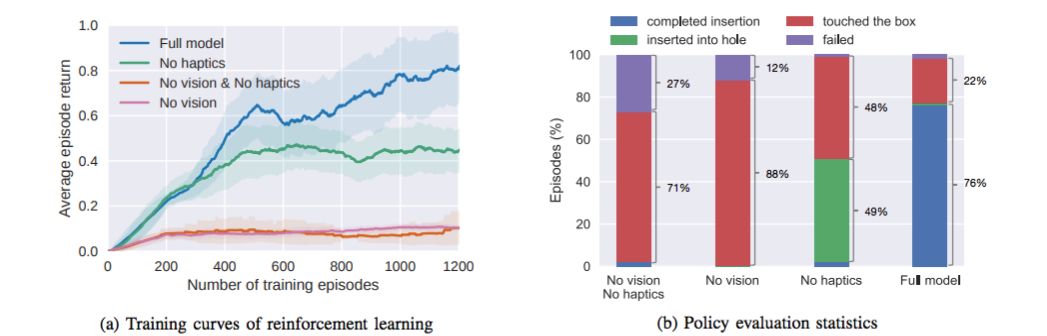


该图为利用自监督进行多模态表征学习的神经网络架构。该网络使用来自三个不同传感器的数据作为模型输入：RGB 图像、力矩数据以及末端执行器的位置、速度。该模型将这些数据编码并融合为多模态表征，基于这些数据可训练模型。

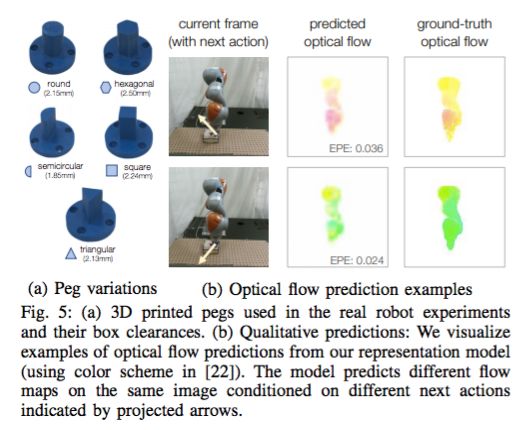
下图展示了该研究的控制器架构，该架构可分为三部分：轨迹生成、阻抗控制和操作空间控制。



下图展示了该模型在模拟环境中的训练。图 a 展示了智能体在结合视触觉或未结合情况下的多条训练曲线，而图b展示了不同反馈模态的任务完成率。



下图展示了在真实环境中的模型评估。

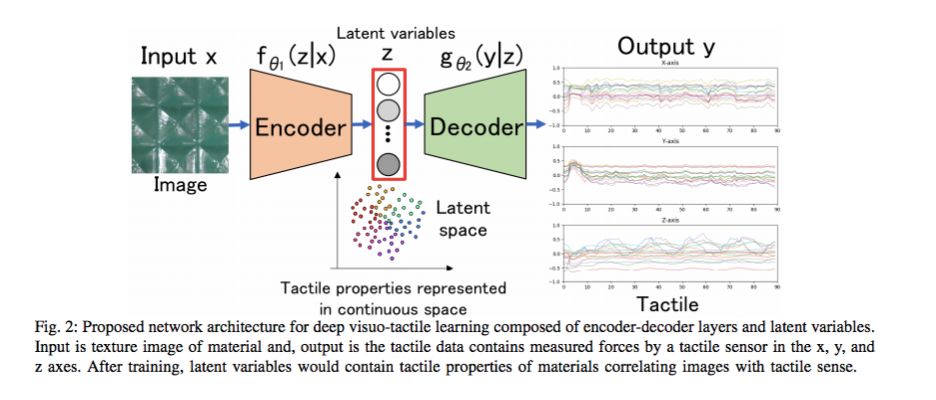


**论文：Deep Visuo-Tactile Learning: Estimation of Tactile Properties from Images**

研究者认为，基于图像对触觉特性的评估会帮助机器人增强对环境的理解，从而在面对具体环境时选择恰当的行为。因此他们提出了一种相关的模型，该方法扩展了编码器-解码器网络，其中主要变量是视觉和触觉特征。

与之前的研究不同，该方法不需要手动标注，仅需要 RGB 图像及对应的触觉感知数据。所有数据都是通过安装在 Sawyer 机器人末端执行器上的网络摄像头和触觉感知器收集的，涉及 25 种不同材料的表面。研究者展示了该模型可以通过评估特征空间，并可以推广到未包含在训练数据中的材料。这表明该模型学会了将图像和重要的触觉特性关联起来。

下图为模型的主要成分，其中，输入为材料的纹理图像，输出为测得的三维触觉数据，处理部分由编码器与解码器等构成。



**论文：Variational End-to-End Navigation and Localization**

研究者扩展了能够理解地图的端到端驾驶网络。他们定义了一个新的变分网络，该网络能够根据环境的原始相机数据和更高级路线图进行学习，进而预测可能的控制指令的完整概率分布，并且在地图内发出确定性控制指令。

此外，研究者根据地图和观察到的视觉道路拓扑之间的对应关系，制定了合适的方案。研究者在真实驾驶数据上评估了该算法，并推断了在不同类型的丰富驾驶场景下命令的稳定性。另外，他们还在一组新的道路和交叉路口上评估了其定位算法，并展示了该模型在无GPS条件下的适用性。

下图描述了模型的工作机理，将原始图像数据与未规划的地图放入卷积通道，并通过全连接层形成GMM。若有已规划地图，则在倒数第二层与上述数据合并处理，形成确定路线。

