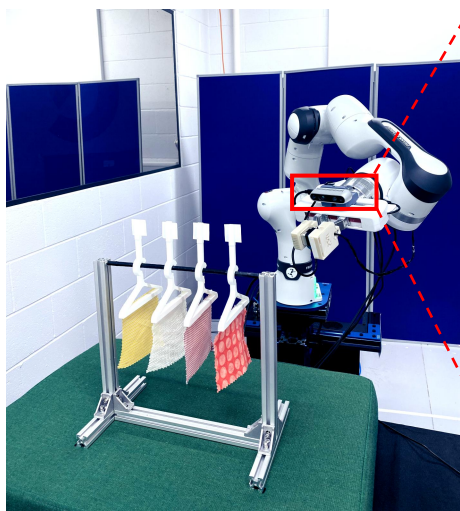
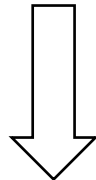


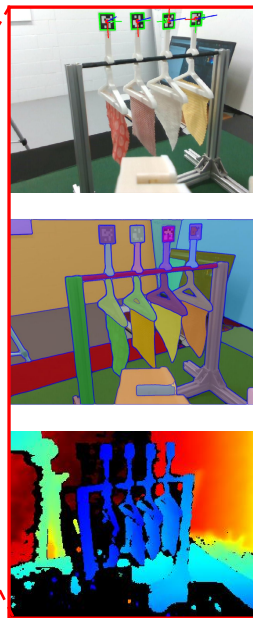
Global Observation



Fabric IDs
Fabric 3D Positions
Fabric RGB Images



Context Memory



Aruco Tags
Detection

Fabric
Segmentation

Depth
Detection

GelSight Images Annotated with
Pressure Values

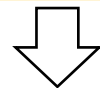
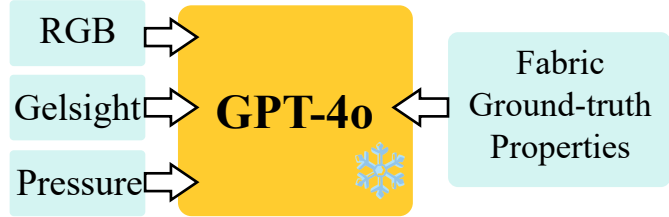


Press Fabric One by One



Sequential Action
Execution

Training & Reasoning



Post Hoc Reasoning

Fabric Property
Reasoning Data

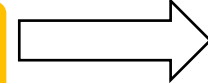
LoRA
Fine-tuning



Ad Hoc Reasoning

Ad Hoc Inference

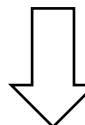
Llama 3.2-Vision



Pick and Deliver



Localization and Action
Execution



Pick me one fabric to...

User