Closing the Perception-Action Loop

Towards General-Purpose Robot Autonomy

Yuke Zhu
Traditional form of **automation**
General-purpose robot hardware
My research goal:

building **robot intelligence** to enrich **human intelligence**
Traditional form of automation

custom-built robots → human expert programming → special-purpose behaviors

New form of automation

general-purpose robots → ? → general-purpose behaviors
Traditional form of **automation**

- **structured** environment
- custom-built robots
- human expert programming
- special-purpose behaviors

New form of **automation**

- **unstructured** environment
- general-purpose robots
- general-purpose behaviors
Traditional form of **automation**

- Custom-built robots
- Human expert programming
- Special-purpose behaviors

New form of **automation**

- General-purpose robots
- Machine learning & perception
- General-purpose behaviors
closing the **perception-action** loop

[Bohg et al. ICRA 2018]

[Levine et al. JMLR 2016]

[Boh et al. ICRA 2018]

closing the **perception-action loop**

with **machine learning & perception in robotics**

Perception is a weak link in the loop.

“Clean” kitchen for state-of-the-art robotics

“Messy” kitchen in the real world
scene understanding

vision & language

affordance reasoning

ICCV’15, CVPR’17a, CVPR’19

CVPR’16, IJCV’17, CVPR’17b

ECCV’14, arXiv’15
two ceramic jars
knives in a holder
green onions sitting on the counter
wooden drawer is closed
a big white bowl

Johnson et al. CVPR’16; Krishna, Zhu, et al. IJCV’17
Q: When was the picture taken?
A: In the daytime.

Q: What color is the countertop?
A: Black.

Q: How many drawer knobs can you see?
A: Seven.
large
on
bowl
next to
bowl
in
holder
knife
in
metal
graspable
has
openable
don
with
counter
black
drawer
knob
closing the perception-action loop

- primitive skills
- sequential tasks
- hierarchical tasks
- general-purpose robot
closing the **perception-action** loop

general-purpose robot

![Diagram](image_url)

primitive skills

sequential tasks

hierarchical tasks

stack

push

grasp

place

pull

RSS’18

CoRL’18

CoRL’18

RSS’18

ICRA’19

closing the perception-action loop

CoRL’18
closing the perception-action loop

primitive skills

sequential tasks

hierarchical tasks

general-purpose robot

“get apple from fridge”

“put bowl in the sink”

ICRA’17

ICCV’17

CVPR’19

“get apple from fridge”

“put bowl in the sink”

ICRA’17

ICCV’17

CVPR’19
closing the **perception-action** loop

**general-purpose robot**

**hierarchical tasks**

**primitive skills**

**sequential tasks**

"prepare dinner"

ICRA'18 CVPR'19

ICCV'15
closing the **perception-action** loop

**Ongoing and Future Work**

- primitive skills
- sequential tasks
- hierarchical tasks

**Humanlike learning robot**

**General-purpose robot**
Part I: Primitive Skills

[Zhu et al., RSS 2018] Reinforcement and Imitation Learning […]


[Zhu*, Fan* et al., CoRL 2018] SURREAL: Open-Source Reinforcement Learning Framework […]
**Primitive Skills** (Reaching, Grasping, Stacking, etc.)

- **perception**
- **action**
- **sensory data**
- **motor command**

Credit: BBC Earth Lab
Primitive Skills  (Reaching, Grasping, Stacking, etc.)

**Challenge:** Raw sensory data are *high-dimensional, noisy, and multimodal.*

**End-to-end learning** of robust representations for control

- sensory data
- motor command
Learning Primitive Skills in Robotics

- reinforcement learning (RL)
- learning through trial-and-error
- requiring a lot of training data

Learning Primitive Skills in Robotics

reinforcement learning (RL)  imitation learning (IL)

learning through trial-and-error  learning from demonstrations

requiring a lot of training data  limited by a suboptimal teacher

[Schaal 1996; Pollard & Hodgins 2002; Abbeel & Ng 2004; Brian et al. 2008; Pastor et al. 2009; Ross et al. 2011; Akgun et al. 2012; Bagnell 2015; Finn et al., 2016; Rahmatizadeh et al., 2017; James et al., 2017; Sermanet et al. 2017; Menda et al. 2017; Le Paine et al. 2018]
Learning Primitive Skills in Robotics

**imitation learning (IL)**
- learning from demonstrations
- limited by a suboptimal teacher

**reinforcement learning (RL)**
- learning through trial-and-error
- requiring a lot of training data

**combined (RL+IL)**
- demonstrations offer guidance
- better performance by trial-and-error

[Price & Boutilier 2003; Bentivegna et al. 2004; Latzke et al. 2007; Conn & Peters 2007; Ross & Bagnell 2014; Kumar et al. 2016; Sun et al. 2017; Vecerik et al. 2007; Rajeswaran et al. 2018; Cheng et al. 2018; Nair et al. 2018; Pfeiffer et al. 2018; Codevilla et al. 2018]
Reinforcement and Imitation Learning

behavioral cloning warm start

inverse reinforcement learning

data augmentation

[Rajeswaran et al. RSS 2018]
Ground-truth states

[Finn et al. ICML 2016]
Short-horizon tasks

[Nair et al. ICRA 2018]
Fixed objects
Reinforcement and Imitation Learning

Task: Car → Red & Plane → Green
Our trained model

Ours

Raw pixel inputs
(RGB camera)

Long-horizon tasks
(each task takes ~1min)

Object variation
(procedural generation)

Prior work

Ground-truth states

Short-horizon tasks

Fixed objects

Zhu, Wang, Merel, Rusu, Erez, Cabi, Tunyasuvunakool, Kramár, Hadsell, de Freitas, Heess, RSS 2018
Reinforcement and Imitation Learning

Effective $\text{RL} + \text{IL} = \text{Algorithm} + \text{Data}$

Task: Car $\rightarrow$ Red & Plane $\rightarrow$ Green

Our trained model

Zhu, Wang, Merel, Rusu, Erez, Cabi, Tunyasuvunakool, Kramár, Hadsell, de Freitas, Heess, RSS 2018
Reinforcement and Imitation Learning: **Algorithm** – Adversarial Learning

**Algorithm**

1. **Policy** \(\pi_\theta\): to fool the discriminator

2. **Discriminator** \(D_\psi\): to tell policy apart from demonstration

---

**Adversarial Learning Objective**

\[
\min_{\theta} \max_{\psi} \mathbb{E}_{\pi_E}[\log D_\psi(s)] + \mathbb{E}_{\pi_\theta}[\log(1 - D_\psi(s))]
\]
Reinforcement and Imitation Learning: **Algorithm** – Adversarial Learning

**IL reward:**

$$r_{IL}(s_t, a_t) = -\log(1 - D_\psi(s_t))$$

**Discriminator objective:**

$$D_\psi$$ predicts 0 if policy and 1 if demo

- **Strong discriminator** predicts 0 if policy
- **Weak discriminator** predicts 1 if demo

**Figure:**
- IL reward
- Discriminator score on policy
- Strong and weak discriminators

---

1Goodfellow et al. 2014; Ho & Ermon, 2016

Zhu, Wang, Merel, Rusu, Erez, Cabi, Tunyasuvunakool, Kramár, Hadsell, de Freitas, Heess, RSS 2018
Reinforcement and Imitation Learning: **Algorithm** – Adversarial Learning

- **Discriminator objective**
  \[ D_\psi(s_t) \]

- Discriminator predicts 0 if policy and 1 if demo

- **IL reward:**
  \[ r_{IL}(s_t, a_t) = -\log(1 - D_\psi(s_t)) \]

- **RL reward:**
  \[ r_{RL}(s_t, a_t) \]

- **RL + IL reward:**
  \[ r_{RL+IL} = \lambda r_{IL}(s_t, a_t) + (1 - \lambda) r_{RL}(s_t, a_t) \]
Reinforcement and Imitation Learning: Algorithm

Input
- 64 x 64 RGB pixel observation
- positions and velocities of arm joints and grippers

Output
- 9-DoF joint velocities at 20Hz

Zh, Wang, Merel, Rusu, Erez, Cabi, Tunyasuvunakool, Kramár, Hadsell, de Freitas, Heess, RSS 2018

Distributed policy gradient method trained on $r_{RL+IL}$
3D motion controller

Collecting human demonstrations

physics engine

Training in simulation

real environment

Running on real robot

Zhu, Wang, Merel, Rusu, Erez, Cabi, Tunyasuvunakool, Kramár, Hadsell, de Freitas, Heess, RSS 2018
block lifting
block stacking
clearing table with blocks

clearing table with a box
pouring liquid
order fulfillment

Zhu, Wang, Merel, Rusu, Erez, Cabi, Tunyasuvunakool, Kramár, Hadsell, de Freitas, Heess, RSS 2018
Reinforcement and Imitation Learning: Algorithm

Zhu, Wang, Merel, Rusu, Erez, Cabi, Tunyasuvunakool, Kramár, Hadsell, de Freitas, Heess, RSS 2018
Reinforcement and Imitation Learning: **Algorithm**

**Result:** Emergent strategies from trial-and-error

- **demonstrated solution**
- **novel solution #1**
  - “grasp two blocks from the top”
- **novel solution #2**
  - “lift both blocks together with three fingers”

*Zhu, Wang, Merel, Rusu, Erez, Cabi, Tunyasuvunakool, Kramár, Hadsell, de Freitas, Heess, RSS 2018*
Reinforcement and Imitation Learning: **Algorithm**

**Result:** Zero-shot sim-to-real policy transfer

visual & dynamics randomization

real-robot deployment

Zhu, Wang, Merel, Rusu, Erez, Cabi, Tunyasuvunakool, Kramár, Hadsell, de Freitas, Heess, RSS 2018
New **algorithm** that learns dexterous *primitive skills*

**Effective RL + IL = Algorithm + Data**

How can we collect **demonstrations** for **diverse skills**?
Reinforcement and Imitation Learning: **Data**

RoboTurk: Crowdsourcing Platform for Large-Scale Demonstration Collection

RoboTurk in action

roboturk.stanford.edu

Mandlekar, Zhu, Garg, Booher, Spero, Tung, Gao, Emmons, Gupta, Orbay, Savarese, Fei-Fei, CoRL 2018
Reinforcement and Imitation Learning: Data

RoboTurk Pilot Dataset

137.5 hours of demonstrations

22 hours of total platform usage

2218 successful demonstrations

teleoperated demonstrations

surreal.stanford.edu
roboturk.stanford.edu

Zhu*, Fan*, Zhu, Liu, Zeng, Gupta, Creus-Costa, Savarese, Fei-Fei, CoRL 2018
Mandlekar, Zhu, Garg, Booher, Spero, Tung, Gao, Emmons, Gupta, Orbay, Savarese, Fei-Fei, CoRL 2018
RoboTurk Pilot Dataset

137.5 hours of demonstrations

22 hours of total platform usage

2218 successful demonstrations

Reinforcement and Imitation Learning: Data

Reinforcement and Imitation Learning
RoboTurk on Physical Robots
RoboTurk for everyone, everywhere
Combining reinforcement and imitation learning to learn primitive skills from raw sensory inputs.

Scaling up demonstration collection with teleoperated crowdsourcing using the RoboTurk platform.
Part II: Sequential Tasks


[Zhu et al., ICCV 2017] Visual Semantic Planning […]

Part I: Primitive Skills    Part II: Sequential Tasks    Part III: Hierarchical Tasks
Sequential Tasks

“put bowl into microwave”

grasp bowl  pull door  place bowl  push door

robot primitive skills

task planning

“put bowl into microwave”

grasp bowl  pull door  place bowl  push door

sequential tasks

Learned plan

Inputs

Task

Put bowl in microwave

Dynamic visual environment

Initial configuration

goal

task

planning
Sequential Tasks

Challenge: very large space of sequential tasks

Sharing and transferring knowledge across tasks

robot primitive skills

millions of tasks, millions of sequences
Prior knowledge is used to learn new tasks faster.

Piaget 1977; Bernstein et al. 1988; Adolph & Berger, 2006; Tuckman & Monetti, 2011; Lake et al., 2016; Tsividis et al., 2017; Dubey et al., 2018
Transfer Learning

source tasks

model

knowledge

Transfer knowledge to similar tasks

target tasks

model

Daumé III, ACL 2007; Bingel & Søgaard, EACL 2017; Bousmalis et al., NIPS 2016; Ganin & Lempitsky, ICML 2015; Ganin et al., JMLR 2016; Kirkpatrick et al., PNAS 2017; Li & Hoiem, PAMI 2017; Luo et al., NIPS 2017; Caruana, ICML 1993; Baxter, ML 1997; Duong et al., ACL 2015; Schaul et al., ICML 2015; Parisotto et al., ICLR 2016; Long et al., NIPS 2017; Hashimoto et al., EMNLP 2017; Rusu et al., CoRL 2017; Lu et al., CVPR 2017.
Transfer Learning

Transfer Learning in Computer Vision
[Zeiler & Fergus 2014; Mahendran & Vedaldi 2015; Huh et al. 2016]

Transfer Learning in Interactive Tasks

pervasive limited
Transfer Learning

Transfer Learning in Computer Vision


Transfer Learning in Interactive Tasks

Rusu et al. 2016; Parisotto et al. 2016; Oh et al. 2017; Barreto et al. 2018

We need a new platform!

Atari games

grid worlds

pervasive

limited
AI2-THOR: A New Platform for Visual AI

ai2thor.allenai.org
AI2-THOR: A New Platform for Visual AI

* developed in Unity 3D game engine

- changing viewpoints
- walking and jumping
- applying forces
- picking & placing
- opening & closing

Zh, Mottaghi, Kolve, Lim, Gupta, Fei-Fei, Farhadi, ICRA, 2017

ai2thor.allenai.org
Visual Task Planning

Interactive Visual Environment

Input

Agent’s view

Task

Putting bowl into microwave

Output

A sequence of high-level commands

<start> → navigate to table → pick up bowl → open microwave → put bowl

Zhu*, Gordon*, Kolve, Fox, Fei-Fei, Gupta, Mottaghi, Farhadi, ICCV 2017
Target-driven Visual Navigation

Input

visual observation

target

Output

target-driven navigation policy*

* Domain adaption with model trained in AI2-THOR

Zhu, Mottaghi, Kolve, Lim, Gupta, Fei-Fei, Farhadi, ICRA, 2017
Visual Task Planning

Goal-directed policy learning

\[ \pi(a|s, g) = \arg \max_a Q(s, a, g) \]

The goal-conditional Bellman equation

\[ Q^*(s, a, g) = \mathbb{E}_\pi [r_g(s, a) + \gamma \max_{a'} Q^*(s', a', g)] \]

Zhu*, Gordon*, Kolve, Fox, Fei-Fei, Gupta, Mottaghi, Farhadi, ICCV 2017
Visual Task Planning

**Key idea:** decoupling environment dynamics and goal specification

\[ Q(s, a, g) \approx \psi(s, a) \top \times \mathbf{w}_g \]

1Dayan 1993; Kulkarni et al., 2016

Zhu*, Gordon*, Kolve, Fox, Fei-Fei, Gupta, Mottaghi, Farhadi, ICCV 2017
Visual Task Planning

**Key idea:** decoupling environment dynamics and goal specification

\[
Q(s_i, a_i, g) = \mathbb{E}\left[ \sum_{i=t}^{\infty} \gamma^{i-t} r_g(s_i, a_i) \right] = \mathbb{E}\left[ \sum_{i=t}^{\infty} \gamma^{i-t} \phi(s_i, a_i)^\top w_g \right] = \mathbb{E}\left[ \sum_{i=t}^{\infty} \gamma^{i-t} \phi(s_i, a_i)^\top \right] w_g = \psi(s_i, a_i)^\top w_g
\]

\(r_g(s, a) = \phi(s, a)^\top w_g\)

reward state-action feature goal embedding

successor feature

\(Q(s, a, g) = 1993; Kulkarni et al., 2016\)

Zhu*, Gordon*, Kolve, Fox, Fei-Fei, Gupta, Mottaghi, Farhadi, ICCV 2017
Visual Task Planning

**Key idea:** decoupling environment dynamics and goal specification

\[
\begin{align*}
    r_{g}(s, a) &= \phi(s, a) \cdot w_{g} \\
    Q(s, a, g) &= \psi(s, a) \cdot w_{g}
\end{align*}
\]

1Dayan 1993; Kulkarni et al., 2016
Visual Task Planning

\[ \pi(a|s, g) = \arg \max_a Q(s, a, g) \]

**Goal-directed Q-value**

\[ Q(s, a, g) \approx \]

**Goal-independent**

\[ \psi(s, a) \]

\[ \times \]

**Goal-specific**

\[ w_g \]

---

Zhu*, Gordon*, Kolve, Fox, Fei-Fei, Gupta, Mottaghi, Farhadi, *ICCV* 2017
Visual Task Planning

\[ \pi(a|s, g) = \arg\max_a Q(s, a, g) \]

\[ Q(s, a, g_0) \approx \psi(s, a)^\top \times w_{g_0} \]

searching for apple

shared across tasks

Zhu*, Gordon*, Kolve, Fox, Fei-Fei, Gupta, Mottaghi, Farhadi, ICCV 2017
Visual Task Planning

$$\pi(a|s, g) = \arg \max_a Q(s, a, g)$$

searching for banana

$$Q(s, a, g_1) \approx \psi(s, a)^\top$$

$$w_{g_1}$$

searching for banana

Zhu*, Gordon*, Kolve, Fox, Fei-Fei, Gupta, Mottaghi, Farhadi, ICCV 2017
Visual Task Planning

$$\pi(a|s, g) = \arg \max_a Q(s, a, g)$$

Putting apple to fridge across tasks

$$Q(s, a, g_2) \approx \psi(s, a)^\top$$

Shared state

Zhu*, Gordon*, Kolve, Fox, Fei-Fei, Gupta, Mottaghi, Farhadi, ICCV 2017
Trained Task: searching for a bowl

Scene → Object → Goal

New task: searching for an egg

\[ \psi(s, a)^\top \]

successor feature

shared

\[ w_g \]

\[ w_g' \]

new goal embedding

Zhu*, Gordon*, Kolve, Fox, Fei-Fei, Gupta, Mottaghi, Farhadi, ICCV 2017
Fast Policy Transfer with New Goal Embedding

**Task:** Search for an egg and put it into the sink

Transfer Learning

Learning from scratch

---

Zhu*, Gordon*, Kolve, Fox, Fei-Fei, Gupta, Mottaghi, Farhadi, *ICCV* 2017
Developed a 3D virtual world (AI2-THOR) to study embodied agents in *interactive visual environments*

**Transfer learning** between *sequential tasks* through the decoupling of environment dynamics and goal specification
Part III: Hierarchical Tasks


[Huang*, Nair*, Xu*, Zhu et al. CVPR 2019] Neural Task Graphs (NTG) [...]

Part I: Primitive Skills  
Part II: Sequential Tasks  
Part III: Hierarchical Tasks
Hierarchical Tasks

“put bowl into microwave”

“prepare dinner”

Complex real-world tasks are hierarchical.
Hierarchical Tasks

Challenge: Task complexity grows exponentially.

“prepare dinner”

Intractable!
Hierarchical Tasks

Challenge: *Task complexity* grows exponentially.

Leveraging the *compositionality* of *hierarchical tasks*
Videos supply massive knowledge to solve new tasks.

Half of YouTube viewers use it to learn how to do things they've never done

Source: The Verge, Pew Research Center
Humans learn efficiently from **video demonstrations**.

**Imitation of Televised Models by Infants**
Andrew N. Meltzoff, *Child Development* 1988

Babies (14-24 months) can learn by imitating demonstrations from the TV screen.
One-Shot Imitation Learning from Videos

Single video demonstration → Meta-learning model → Policy for the demonstrated task

Xu*, Nair*, Zhu, Gao, Garg, Fei-Fei, Savarese. ICRA 2018
One-Shot Imitation Learning from Videos

a lot of training videos (seen tasks)

supervision

meta-learning model

policy for the demonstrated task

Xu*, Nair*, Zhu, Gao, Garg, Fei-Fei, Savarese. ICRA 2018
One-Shot Imitation Learning from Videos

Xu*, Nair*, Zhu, Gao, Garg, Fei-Fei, Savarese. ICRA 2018
One-Shot Imitation Learning from Videos

modeling demonstration as a **flat sequence**

modeling demonstration as a **compositional structure**

[Duan et al. 17; Finn et al. 2017; Wang et al. 2017; Yu et al. 2018]

Xu*, Nair*, Zhu, Gao, Garg, Fei-Fei, Savarese. ICRA 2018
Neural Task Programming (NTP): Hierarchical Policy Learning as Neural Program Induction
One-Shot Imitation Learning from Videos: **Neural Task Programming (NTP)**

- **Observation**
  - Demonstration policy
  - Next program
  - Current program: `pick_place(blue, green)`

- **End-to-End Neural Network (LSTM)**
- **Meta-Learning Model**
- **Robot API**
  - Completed Tasks

---

Xu*, Nair*, Zhu, Gao, Garg, Fei-Fei, Savarese. *ICRA 2018*
One-Shot Imitation Learning from Videos: **Neural Task Programming (NTP)**

![Diagram of NTP](image)

**Demonstration**

- Task 1 Final State
- Task 2 Final State

**Observation**

- Task 1
- Task 2

**Current Program**

- pick_place(blue, green)

**End-to-End Neural Network (LSTM)**

**Next Program**

- pick(blue)

**Training Supervision**

- Video Demonstration
- Hierarchical Program Trace

---

Xu*, Nair*, Zhu, Gao, Garg, Fei-Fei, Savarese. *ICRA 2018*
One-Shot Imitation Learning from Videos: **Neural Task Programming (NTP)**

**Qualitative**

**Quantitative**

Better generalization with less training data than flat baselines

Unseen task success rate

<table>
<thead>
<tr>
<th></th>
<th>NTP (Ours)</th>
<th>N/A</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Object Sorting

Autonomous Execution

Demo

Xu*, Nair*, Zhu, Gao, Garg, Fei-Fei, Savarese. *ICRA 2018*
One-Shot Imitation Learning from Videos: **Neural Task Programming (NTP)**

- Demonstration
- Meta-learning model
- End-to-end neural network (LSTM)
- Policy
- Compositional model prior

---

Xu*, Nair*, Zhu, Gao, Garg, Fei-Fei, Savarese. ICRA 2018
One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**

Diagram:
- Demonstration
- Neural Task Graph Generator
- Neural Task Graph
- Meta-learning model
- Task Graph Executor
- Policy
- Observation

Huang*, Nair*, Xu*, Zhu, Garg, Fei-Fei, Savarese, Niebles. *CVPR 2019*
One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**

**Task Graph**

- Node: place(red)
- Node: place(green)
- Node: pick(red)
- Node: pick(green)
- Node: pick(orange)

**Conjugate Task Graph**

- Node: place(green)
- Node: pick(green)
- Node: pick(red)
- Node: pick(orange)

**Valid states**

- States: {last, ..., first}

**Nodes** | **States** | **Actions**
---|---|---
finite | infinite | finite

Huang*, Nair*, Xu*, Zhu, Garg, Fei-Fei, Savarese, Niebles. *CVPR 2019*
One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**

**Diagram:**
- **Current Observation**: A scene with objects labeled with colors and letters.
- **Node Localizer**: Connects the current observation to the node.
- **Edge Classifier**: Determines the next action.
- **Selected Node**: Represents the next action to be performed.
- **Place**: Action to place an object.
- **Pick**: Action to pick an object.

Actions:
- **Pick (red)**
- **Pick (green)**
- **Pick (orange)**
- **Place (red)**

References:
Huang*, Nair*, Xu*, Zhu, Garg, Fei-Fei, Savarese, Niebles. *CVPR 2019*
One-Shot Imitation Learning from Videos: Neural Task Graphs (NTG)

Training supervision

video demonstration

action sequence

Demonstration

Task Graph Generator

Neural Task Graph

Task Graph Executor

policy

observation

Huang*, Nair*, Xu*, Zhu, Garg, Fei-Fei, Savarese, Niebles. CVPR 2019
One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**

**Qualitative**

**Autonomous Execution**

**Quantitative**

![Graph showing success rate vs. number of training tasks]

Weaker supervision, less training data, and better generalization

**Recovery from Intermediate Failures**

**Weaker supervision, less training data, and better generalization**

Huang*, Nair*, Xu*, Zhu, Garg, Fei-Fei, Savarese, Niebles. *CVPR 2019*
Extracting knowledge about the compositional structure of hierarchical tasks from video demonstrations

Meta-learning models with compositional priors generalize better than black-box models
closing the **perception-action** loop

- **primitive skills**
- **sequential tasks**
- **hierarchical tasks**
- **general-purpose robot**
<table>
<thead>
<tr>
<th>Closing the Perception-Action Loop</th>
<th>Perception Modality</th>
<th>Action Abstraction</th>
<th>Learning Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>primitive skills</strong> [RSS'18, CoRL'18a, CoRL'18b]</td>
<td>real-world sensory data</td>
<td>joint torque</td>
<td>reinforcement &amp; imitation learning</td>
</tr>
<tr>
<td><strong>sequential tasks</strong> [ICRA'17, ICCV'17]</td>
<td>interactive visual environment</td>
<td>high-level command</td>
<td>transfer learning</td>
</tr>
<tr>
<td><strong>hierarchical tasks</strong> [ICRA'18, CVPR'19]</td>
<td>unstructured video data</td>
<td>task structure</td>
<td>meta-learning</td>
</tr>
</tbody>
</table>
closing the perception-action loop

Ongoing and Future Work

primitive skills
sequential tasks
hierarchical tasks

humanlike learning robot
genral-purpose robot
Future Direction: Multimodal Perception Beyond Vision

Learning coherent representations of multimodal information for control

combining vision and force for manipulation

Future Direction: Learning Knowledge of the World from Interaction

**semantic**

[Zhu et al. CVPR’16; Krishna, Zhu et al. IJCV’17; Xu, Zhu et al. CVPR’17; Zhu et al. CVPR’17]

**geometric**

[Chen, Xu, Zhu et al. CVPR’19]

**physical**

[Zhu et al. ECCV’14; Fang, Zhu et al. RSS’18]

**causal**

[ongoing work]

"To accelerate or to brake?"
Future Direction: Integrating Perception and Knowledge for Autonomy

Data-driven + Model-driven Methods

ongoing work

broader collaboration