USC University of Southern California

U.S. National Science Foundation

NSF: Safe Learning-Enabled Systems

A Neurosymbolic Approach for Safe Multi-Agent Systems

The Team

Lars Lindemann

FOLLOWING

Assistant Professor of Computer Science, University of Southern California
Verified email at usc.edu - Homepage

Systems and Control Theory · Formal Methods · Machine Learning · Autonomous Systems

Jyotirmoy V. Deshmukh

Associate Professor, University of Southern California
Verified email at usc.edu - Homepage

Cyberphysical systems · Formal/Statistical Verification · Temporal logic · AI safety · Reinforcement learning

Mahdi Soltanolkotabi

University of Southern California
Verified email at usc.edu - Homepage

Theoretical Foundations of AI · AI for science/med · (non)convex optimization · probability · statistics

Gaurav Sukhatme

Professor, Departments of CS and ECE, USC
Verified email at usc.edu - Homepage

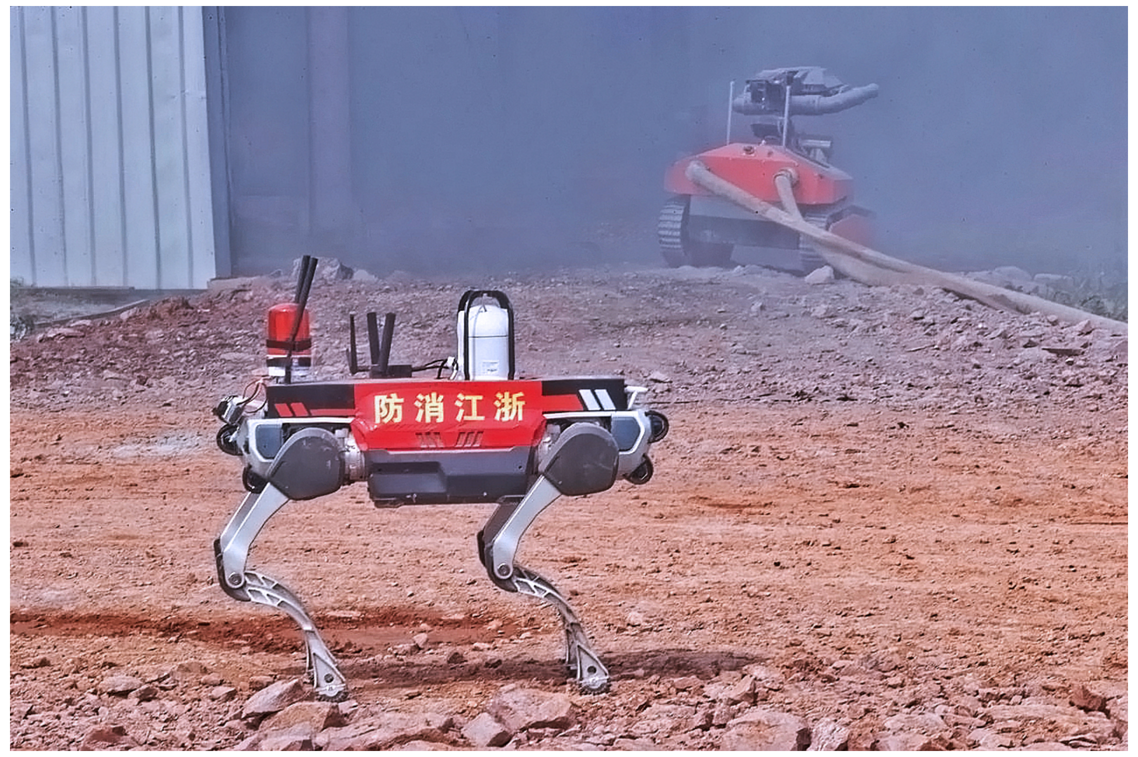
Robotics · Artificial Intelligence · Robot Networks · Motion Planning · Machine Learning

Learning-enabled multi-agent systems (LE-MAS)

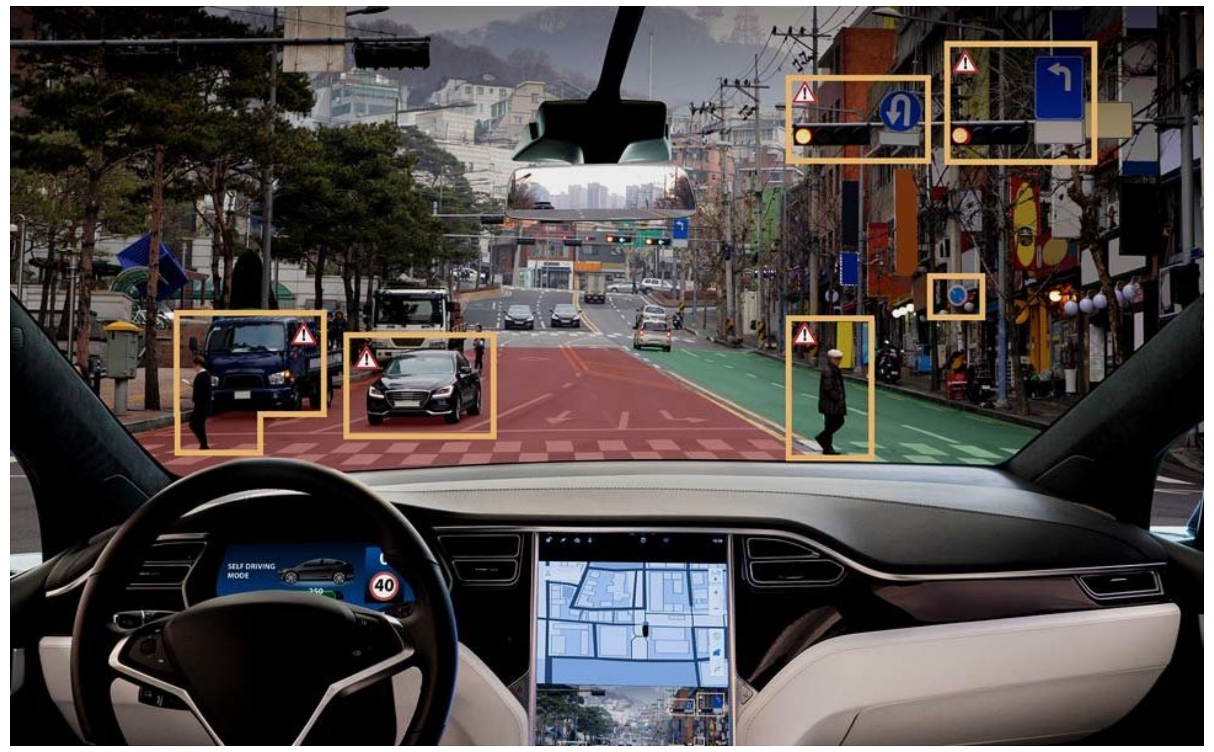
A learning-enabled multi-agents system (LE-MAS) is a network of intelligent agents that utilize learning-enabled components (LECs).

Wildfire prevention

Earthquake recovery



Autonomous driving



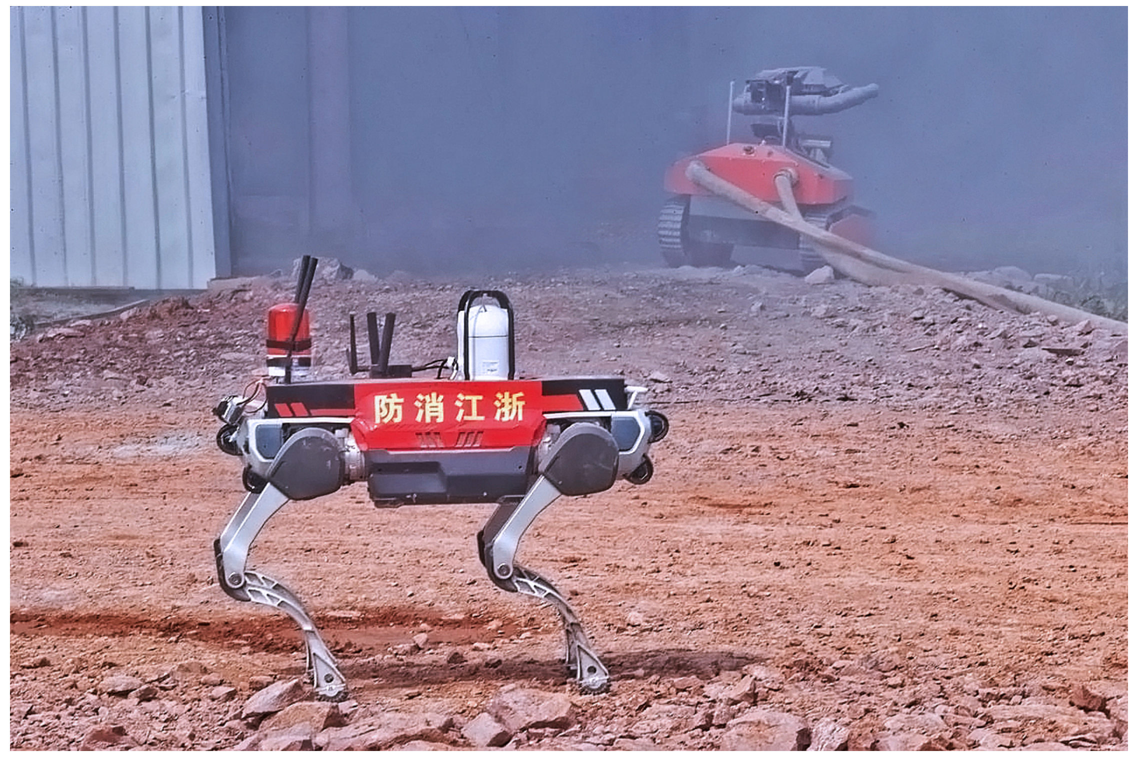
Warehouses

Learning-enabled multi-agent systems (LE-MAS)

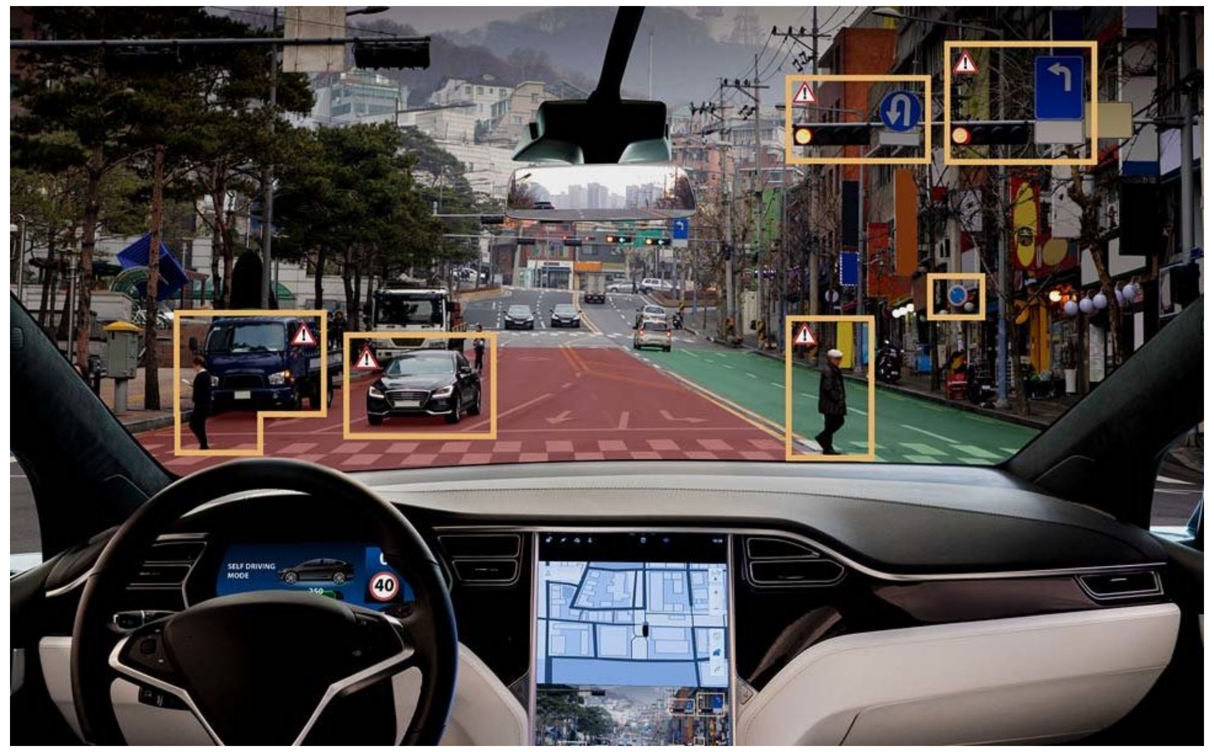
A learning-enabled multi-agents system (LE-MAS) is a network of intelligent agents that utilize learning-enabled components (LECs).

Wildfire prevention

Earthquake recovery



Autonomous driving



Warehouses

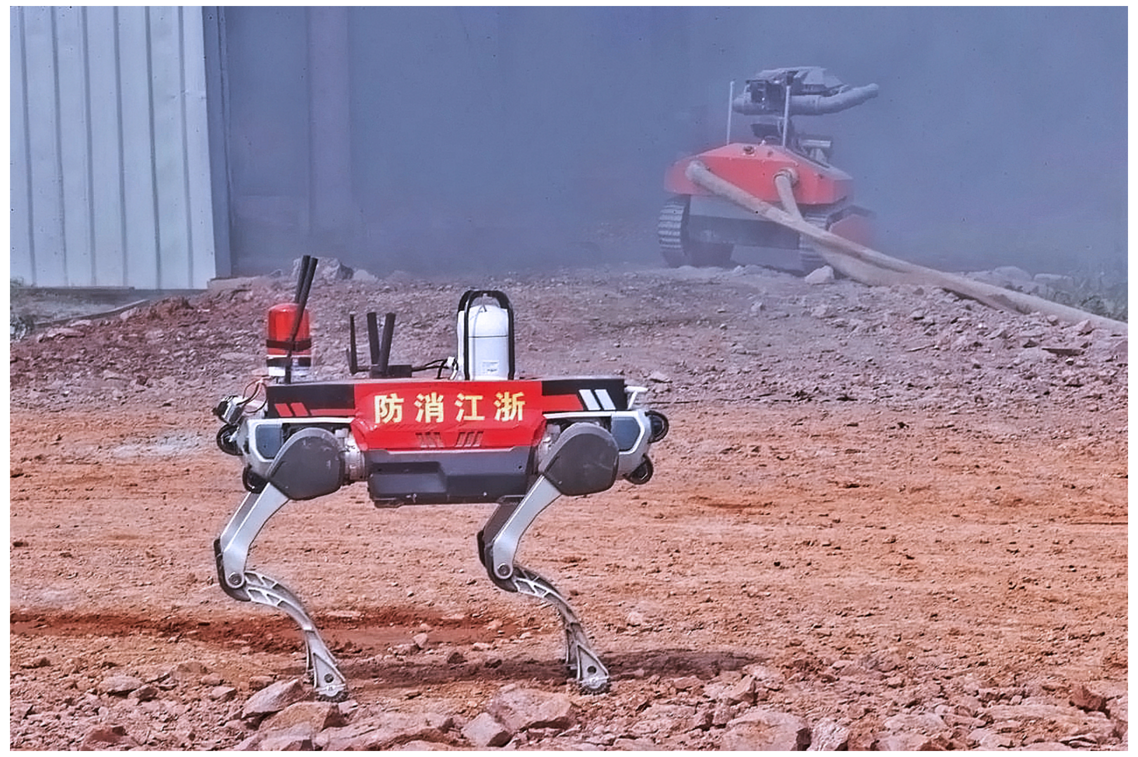
Observations:

Learning-enabled multi-agent systems (LE-MAS)

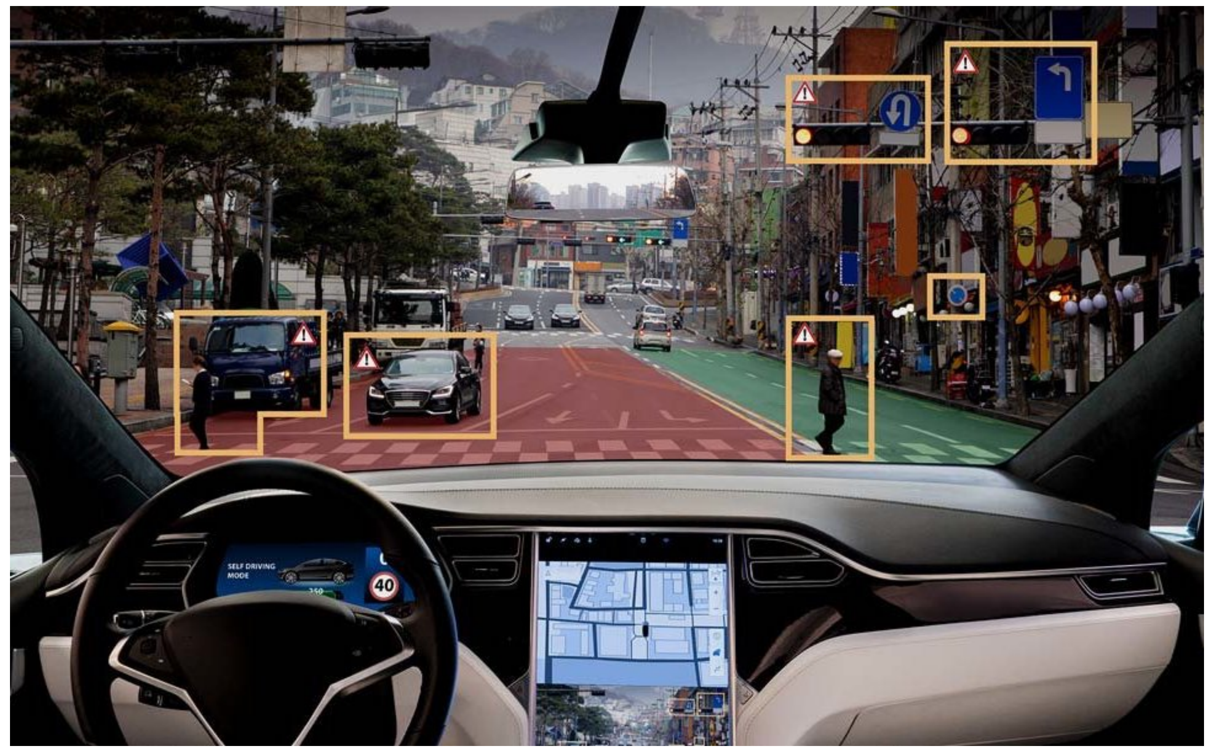
A learning-enabled multi-agents system (LE-MAS) is a network of intelligent agents that utilize learning-enabled components (LECs).

Wildfire prevention

Earthquake recovery



Autonomous driving



Warehouses

Observations:

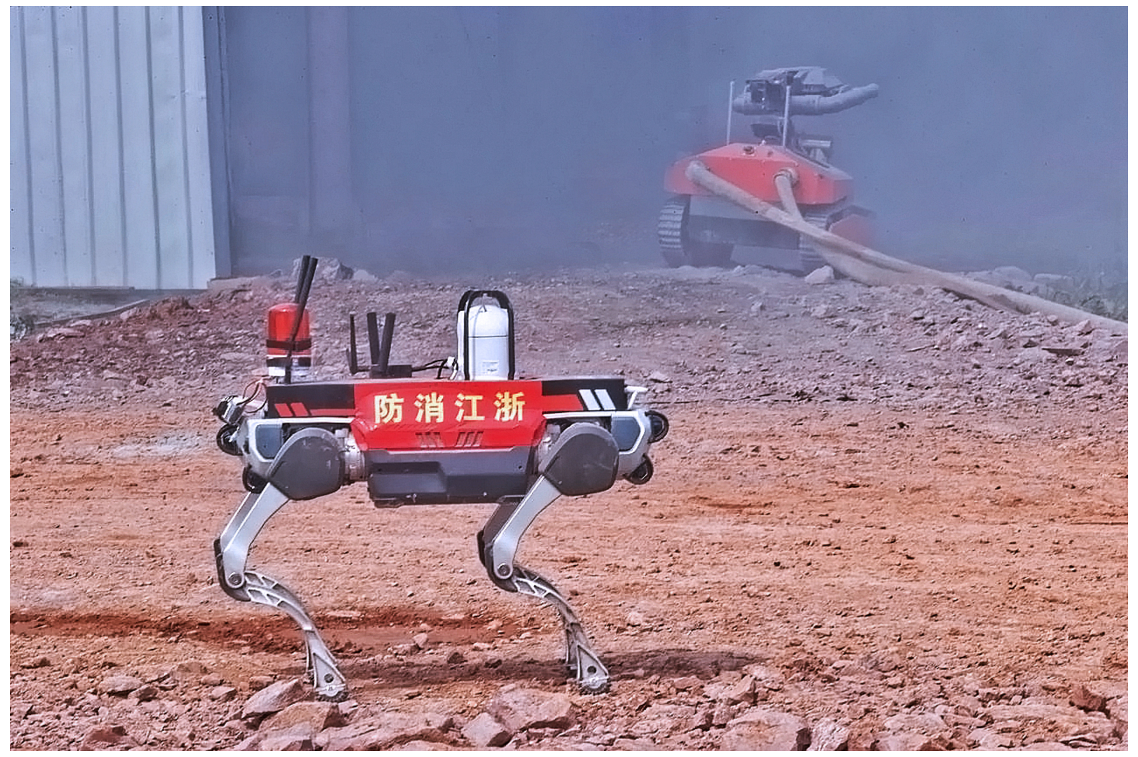
- Multi-agent systems are typically controlled by human operators!

Learning-enabled multi-agent systems (LE-MAS)

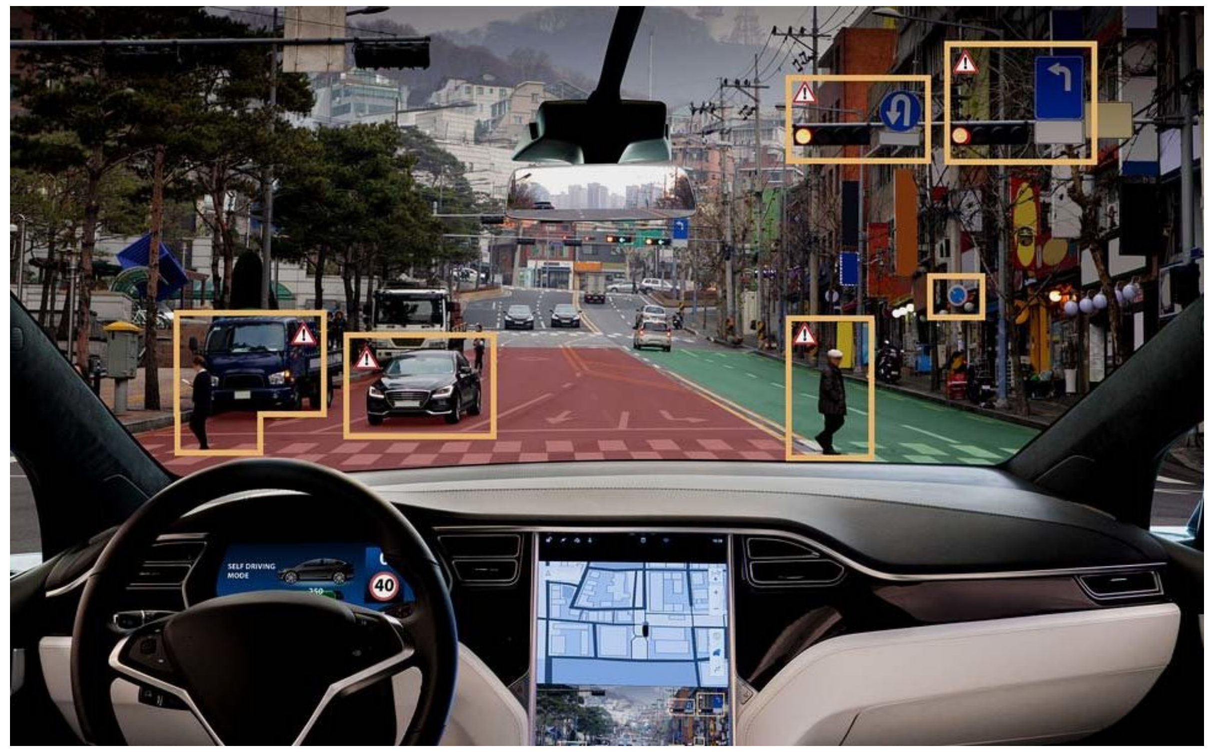
A learning-enabled multi-agents system (LE-MAS) is a network of intelligent agents that utilize learning-enabled components (LECs).

Wildfire prevention

Earthquake recovery



Autonomous driving



Warehouses

Observations:

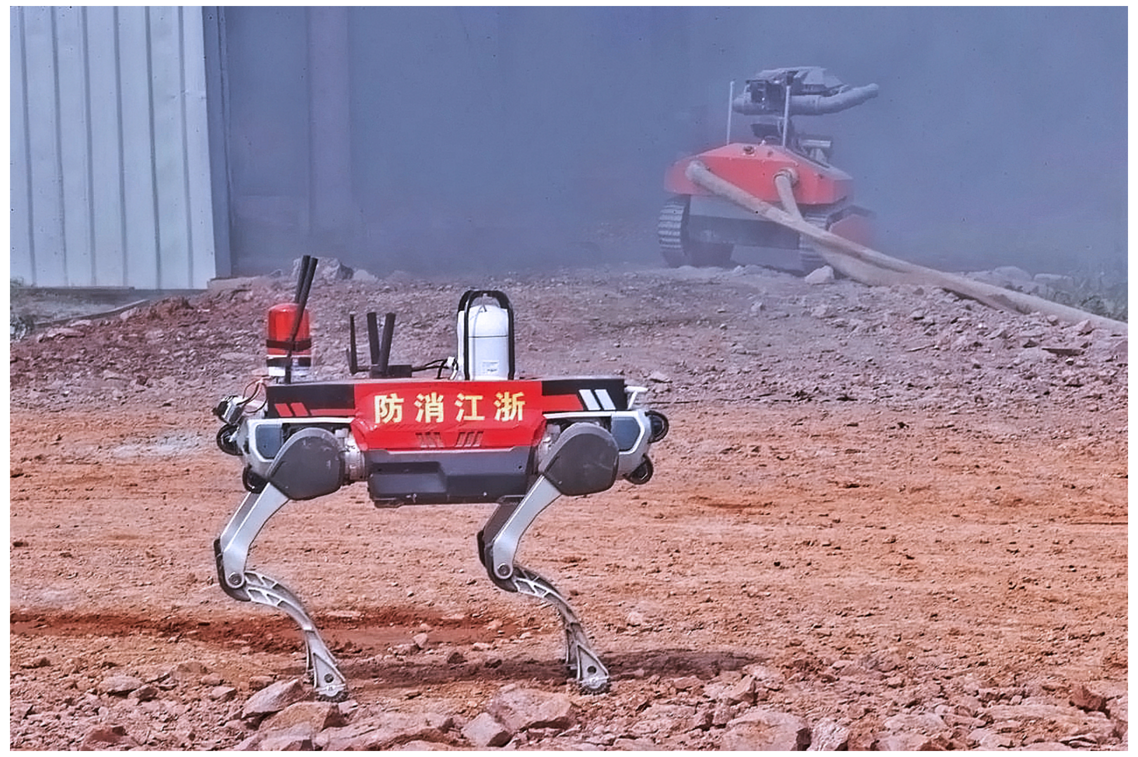
- Multi-agent systems are typically controlled by human operators!
- LE-MASs problems are challenging: (1) effect of uncertainty of LECs on network structure, (2) data dependencies between agents, and (3) high-dimensional problem space.

Learning-enabled multi-agent systems (LE-MAS)

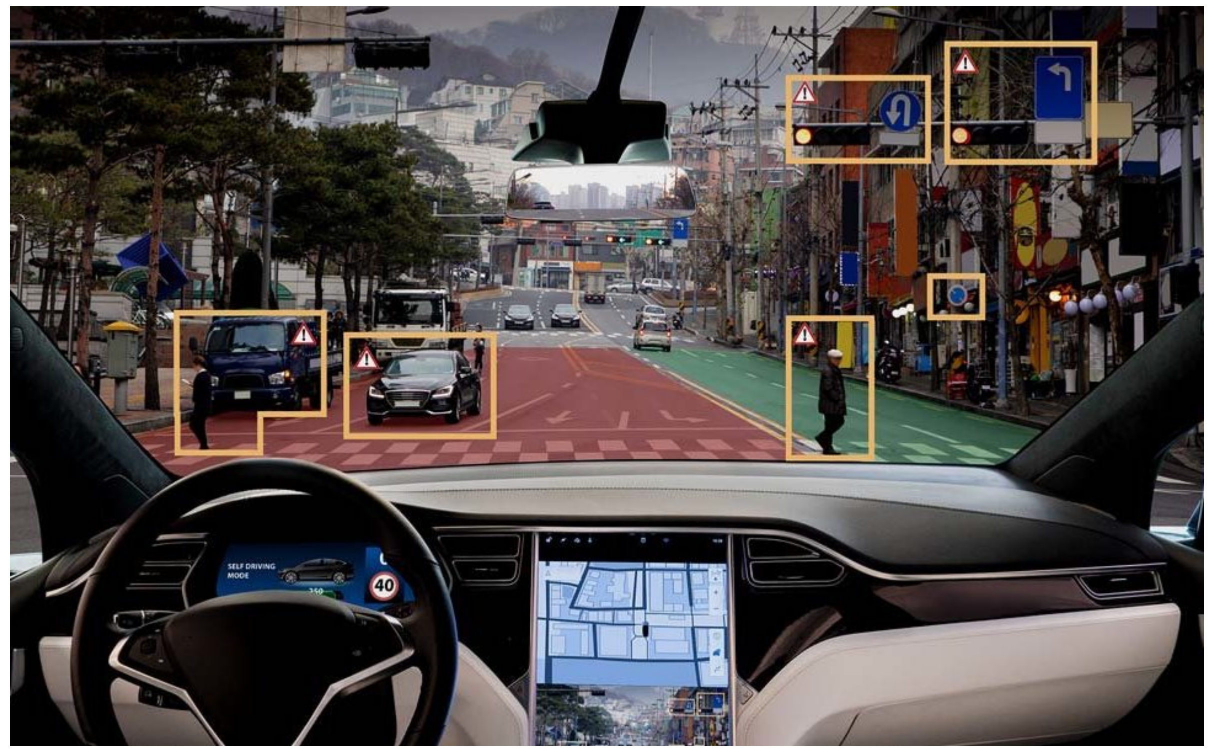
A learning-enabled multi-agents system (LE-MAS) is a network of intelligent agents that utilize learning-enabled components (LECs).

Wildfire prevention

Earthquake recovery



Autonomous driving



Warehouses

Observations:

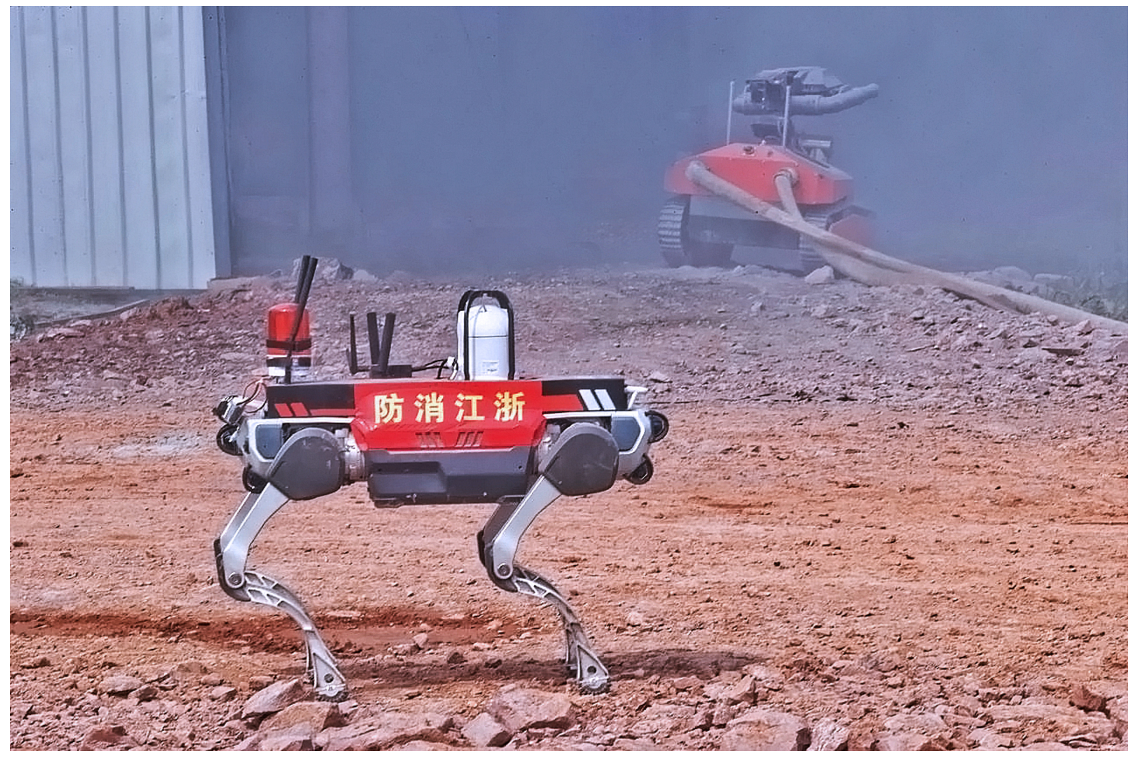
- Multi-agent systems are typically controlled by human operators!
- LE-MASs problems are challenging: (1) effect of uncertainty of LECs on network structure, (2) data dependencies between agents, and (3) high-dimensional problem space.
- Research gap: while LE-SASs are fairly well understood, it is unclear how to design safe LE-MASs.

Learning-enabled multi-agent systems (LE-MAS)

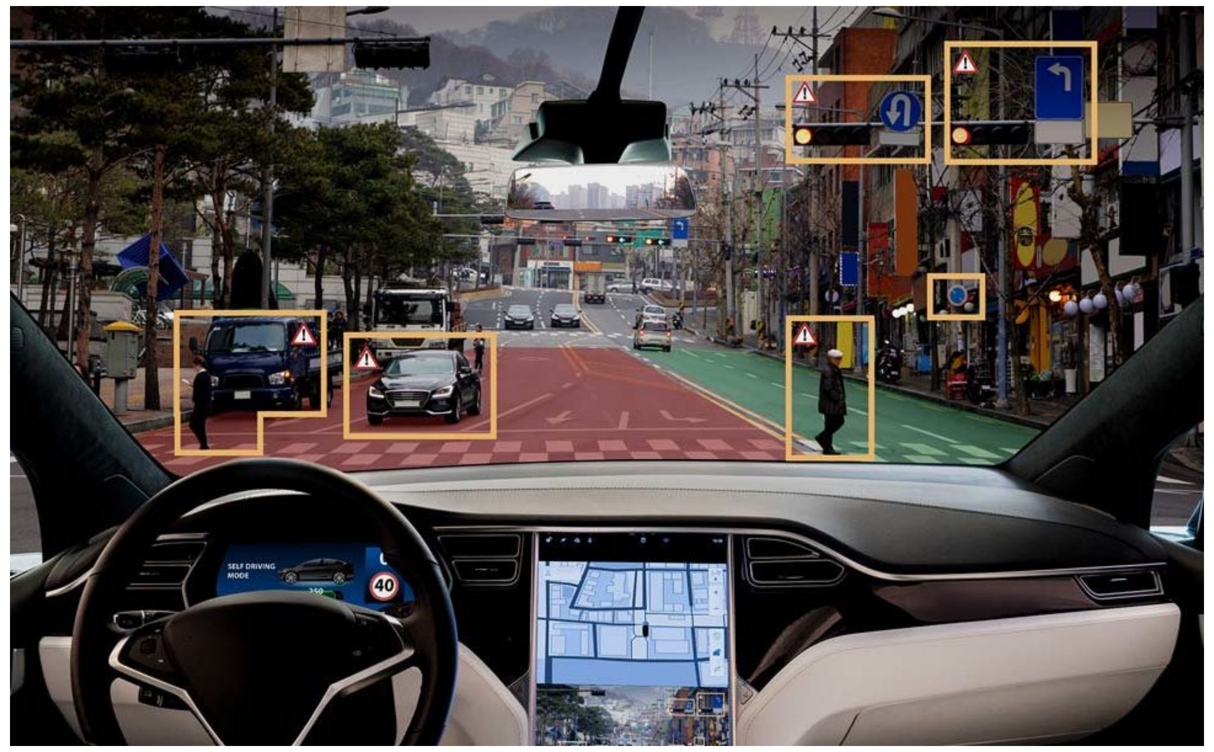
A learning-enabled multi-agents system (LE-MAS) is a network of intelligent agents that utilize learning-enabled components (LECs).

Wildfire prevention

Earthquake recovery



Autonomous driving



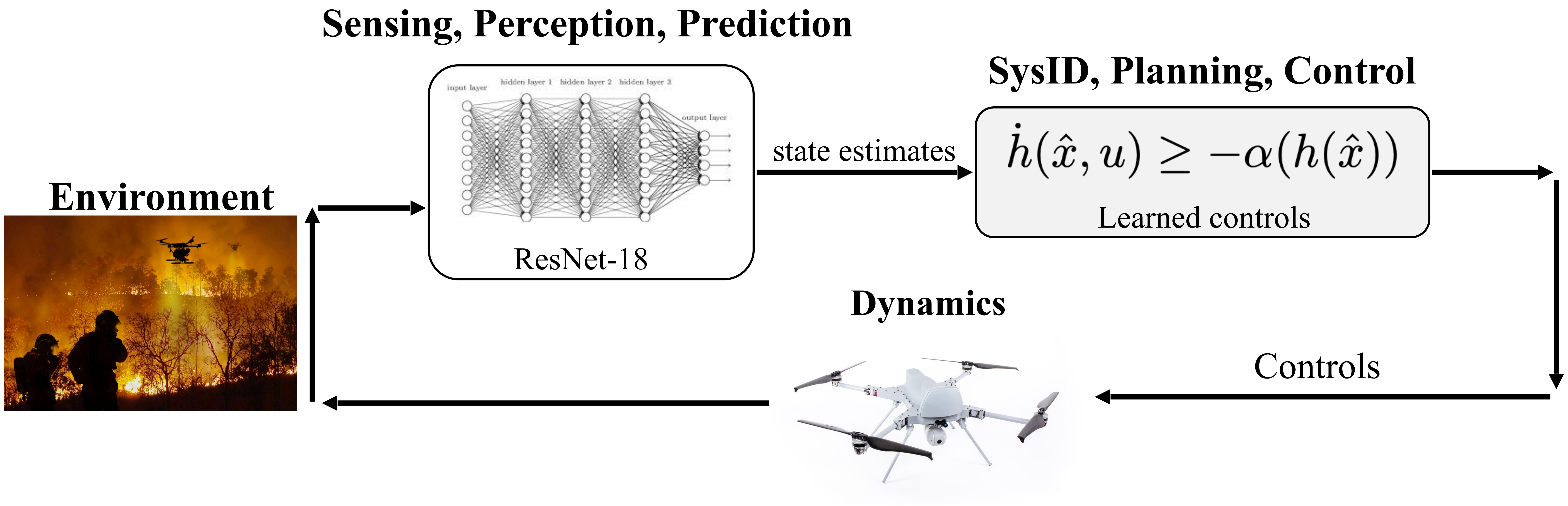
Warehouses

out-of-distribution situations
more likely and severe

Observations:

- Multi-agent systems are typically controlled by human operators!
- LE-MASs problems are challenging: (1) effect of uncertainty of LECs on network structure, (2) data dependencies between agents, and (3) high-dimensional problem space.
- Research gap: while LE-SASs are fairly well understood, it is unclear how to design safe LE-MASs.

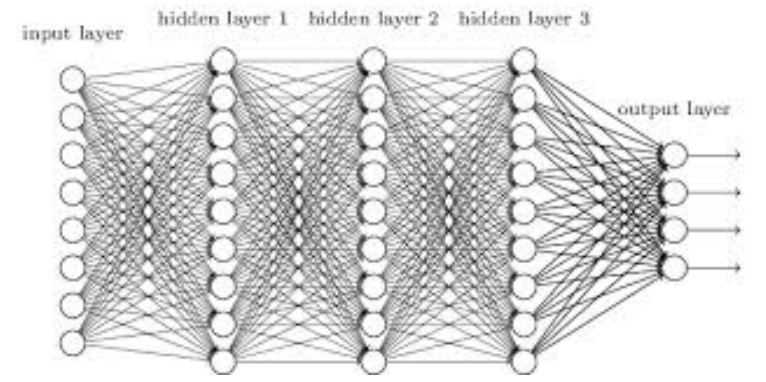
Learning-enabled components



Learning-enabled components

Sensing, Perception, Prediction

Environment



ResNet-18

state estimates

SysID, Planning, Control

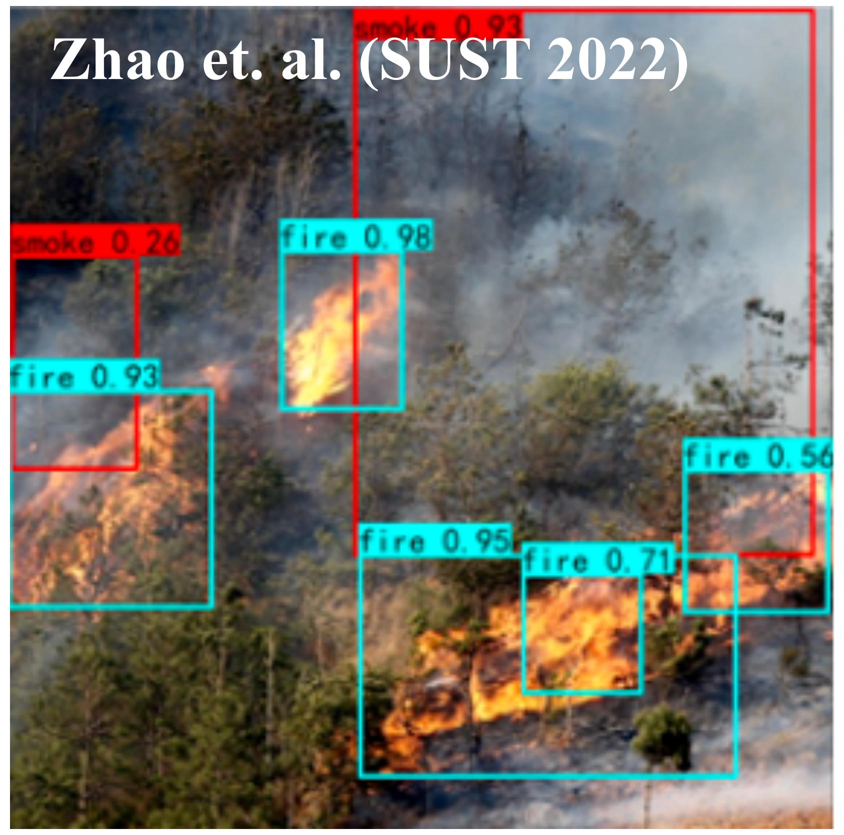
$$\dot{h}(\hat{x}, u) \geq -\alpha(h(\hat{x}))$$

Learned controls

Controls

Dynamics

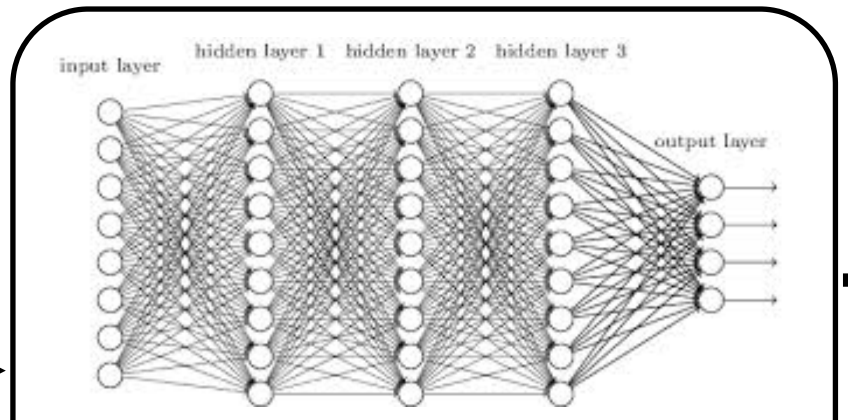
Learning-enabled components:



Object detection

Learning-enabled components

Sensing, Perception, Prediction



ResNet-18

state estimates

SysID, Planning, Control

$$\dot{h}(\hat{x}, u) \geq -\alpha(h(\hat{x}))$$

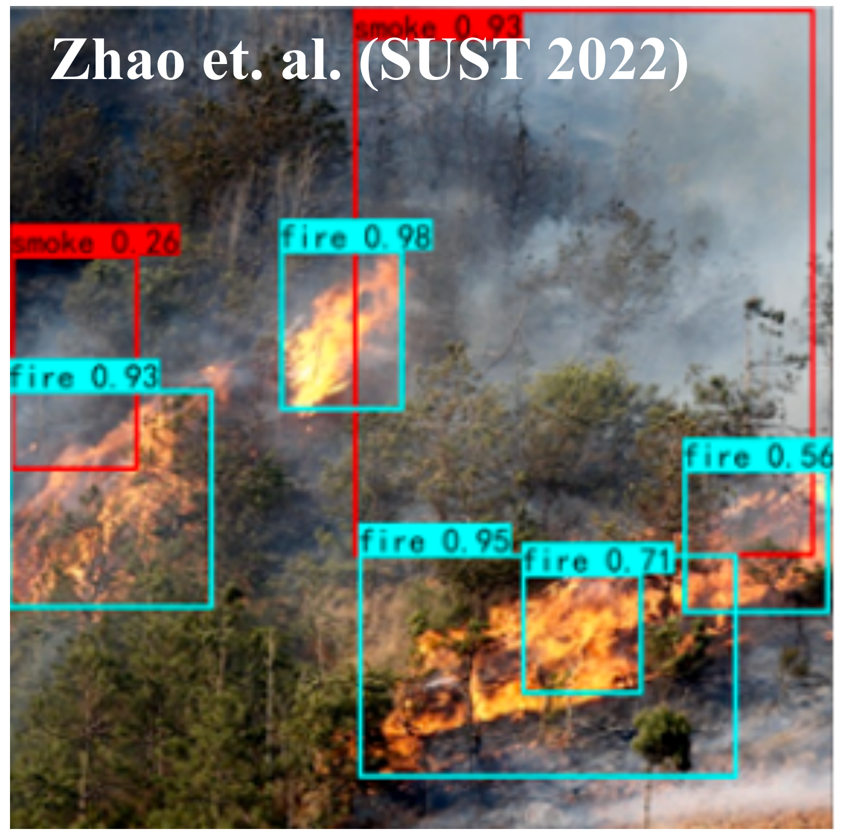
Learned controls

Environment

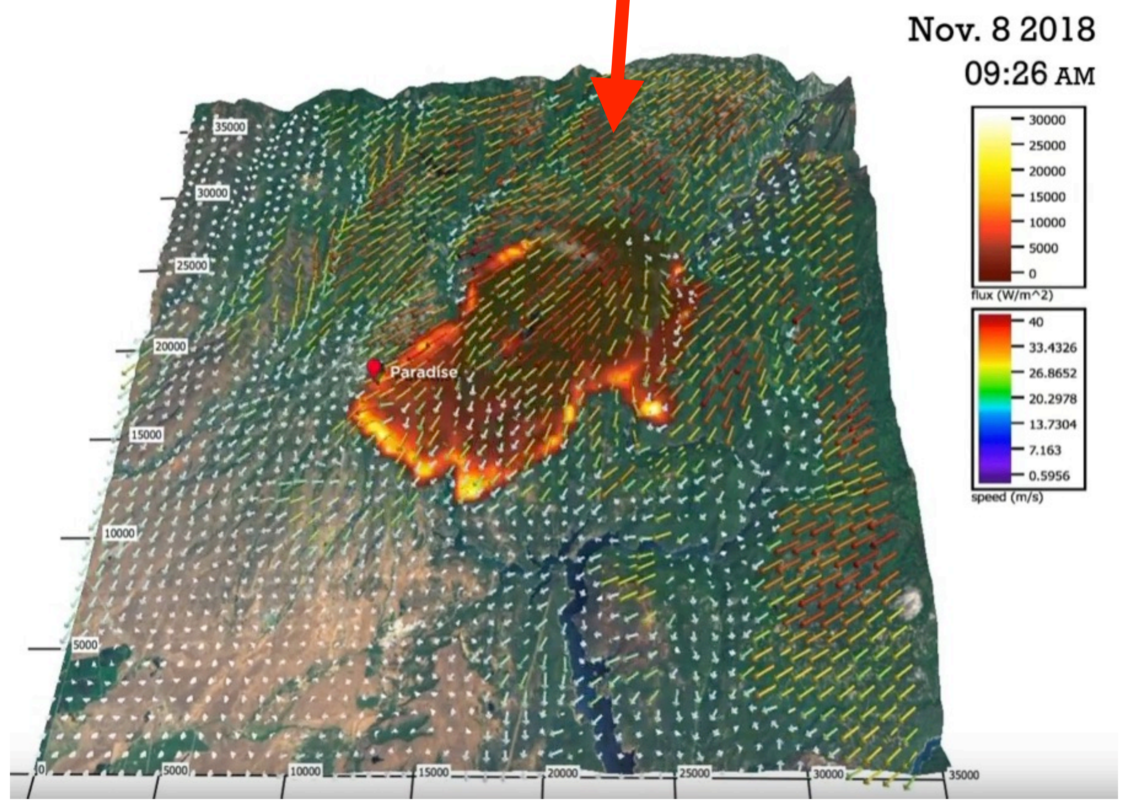
Dynamics

Controls

Learning-enabled components:

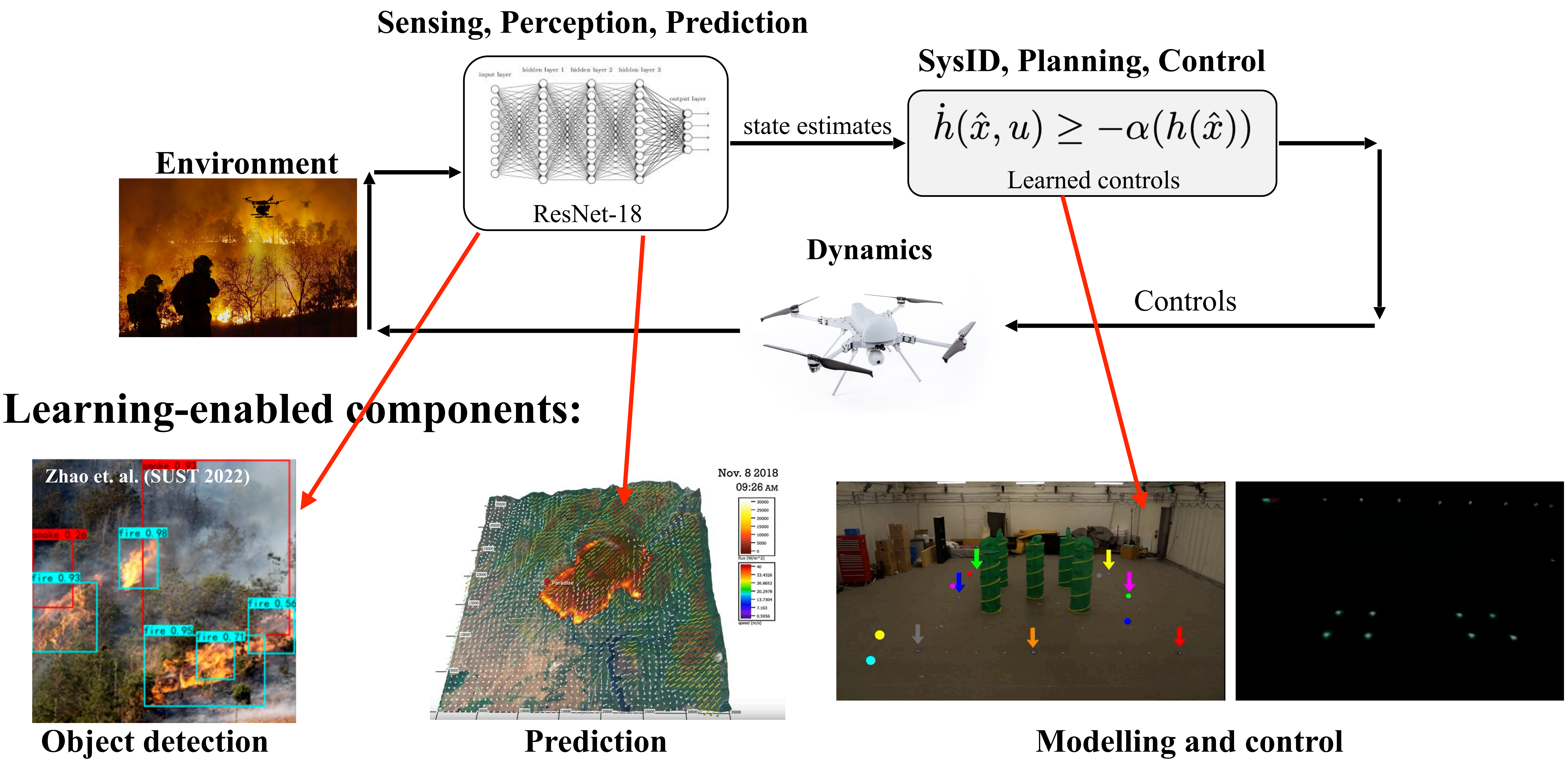


Object detection



Prediction

Learning-enabled components



Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

(1) Formal specifications to reason over LE-MASs with multiple network topologies

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

(1) Formal specifications to reason over LE-MASs with multiple network topologies

via STL-GO

"redundant agent-to-agent perception, robust wildfire detection, avoiding strong wind gusts, …"

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

LE-MASs challenges: data dependencies among agents, more prone to out-of-distribution behavior

Open problems and proposed work:

(1) Formal specifications to reason over LE-MASs with multiple network topologies

via STL-GO

“redundant agent-to-agent perception, robust wildfire detection, avoiding strong wind gusts, …”

(2) Statistical neurosymbolic verification

FBI working to piece together drone that damaged super scooper while battling Palisades Fire

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

LE-MASs challenges: data dependencies among agents, more prone to out-of-distribution behavior

Open problems and proposed work:

(1) Formal specifications to reason over LE-MASs with multiple network topologies

via STL-GO

"redundant agent-to-agent perception, robust wildfire detection, avoiding strong wind gusts, …"

(2) Statistical neurosymbolic verification

FBI working to piece together drone that damaged super scooper while battling Palisades Fire

Symbolic verification does not scale for LE-MASs!

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

LE-MASs challenges: data dependencies among agents, more prone to out-of-distribution behavior

Open problems and proposed work:

(1) Formal specifications to reason over LE-MASs with multiple network topologies

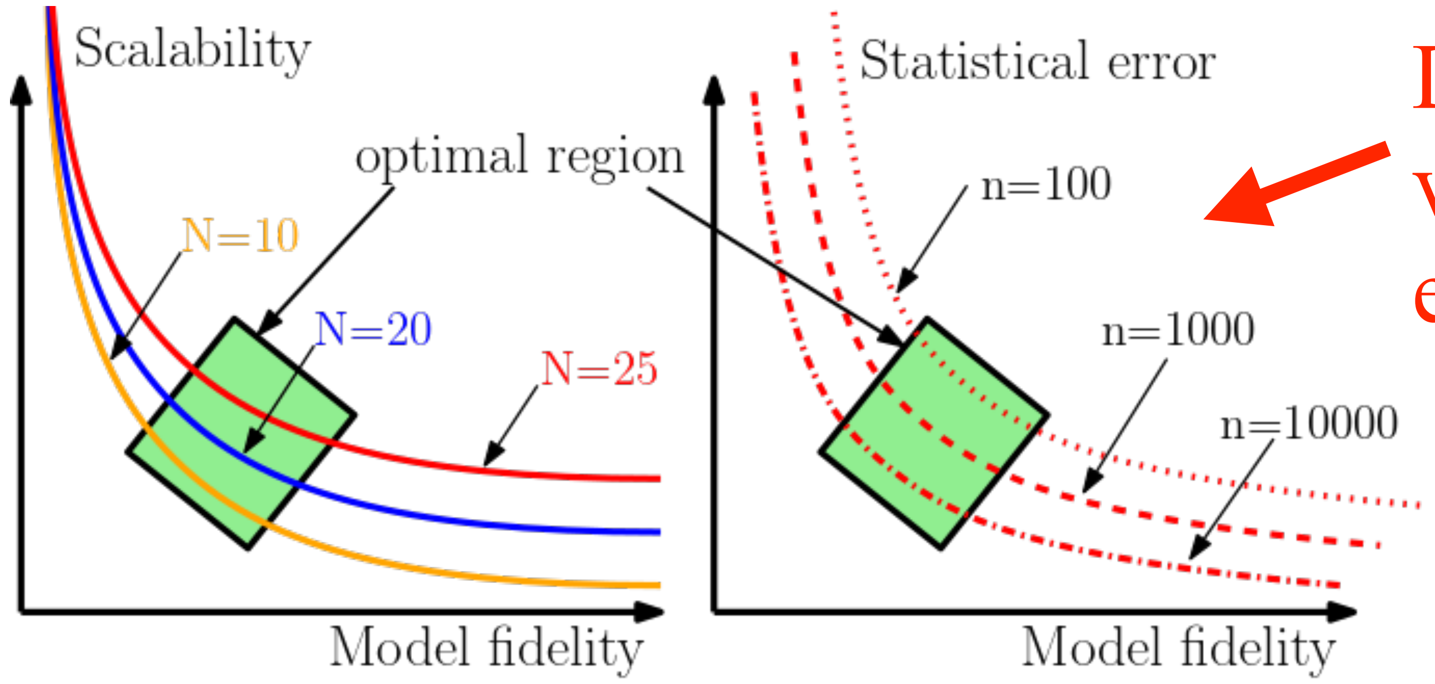
via STL-GO

"redundant agent-to-agent perception, robust wildfire detection, avoiding strong wind gusts, …"

(2) Statistical neurosymbolic verification

FBI working to piece together drone that damaged super scooper while battling Palisades Fire

Symbolic verification does not scale for LE-MASs!



Distributed surrogate model verification and statistical error quantification!

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

(3) Reinforcement learning for STL-GO

How to avoid reward hacking in LE-MASs?

Statistical neurosymbolic reasoning for LE-MASs

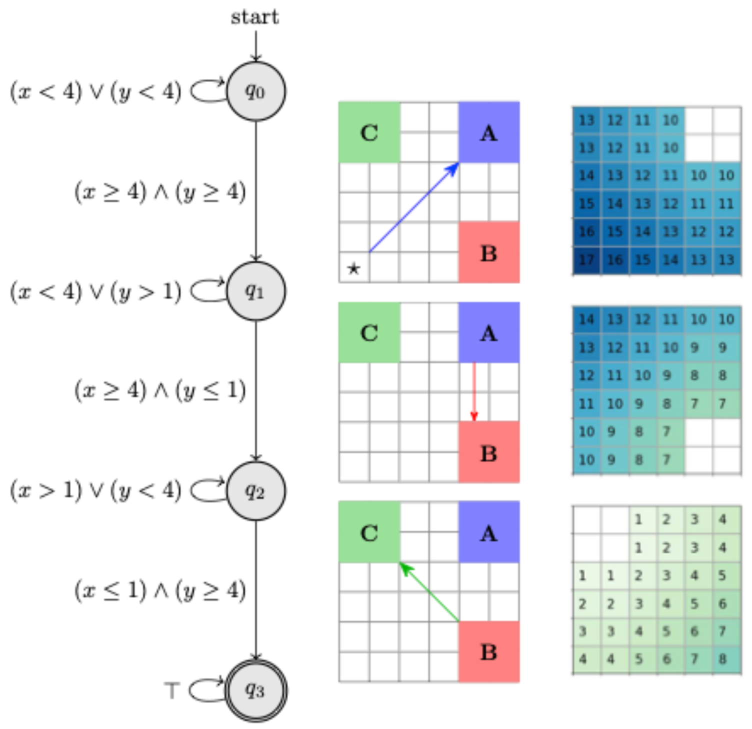
Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

(3) Reinforcement learning for STL-GO

How to avoid reward hacking in LE-MASs?

- Decentralized MARL with STL-GO rewards
- Sample efficiency via symbolic automata



Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

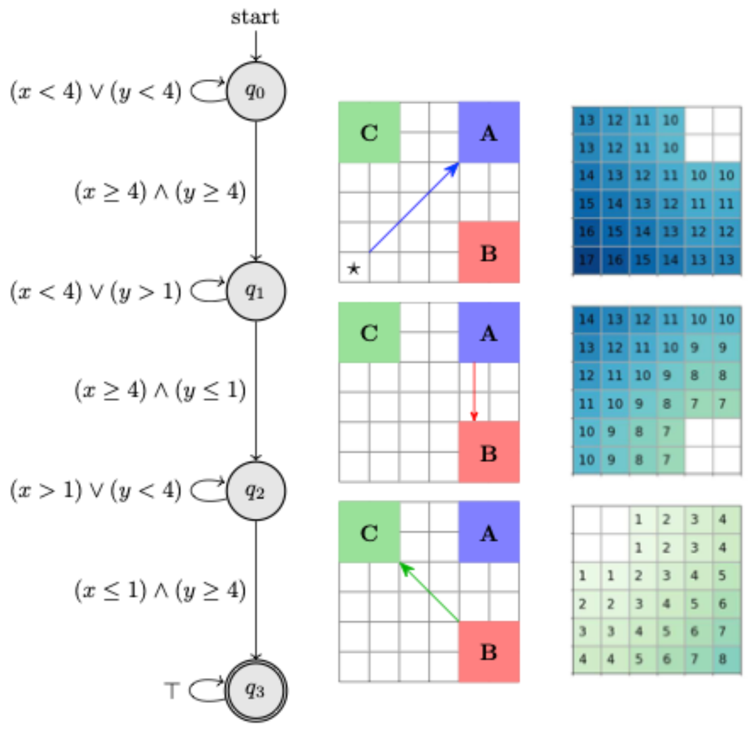
Open problems and proposed work:

Sim2real gap, data corruption, …

(3) Reinforcement learning for STL-GO

(4) Robust LEC training under distribution shift

How to avoid reward hacking in LE-MASs?



- Decentralized MARL with STL-GO rewards
- Sample efficiency via symbolic automata

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

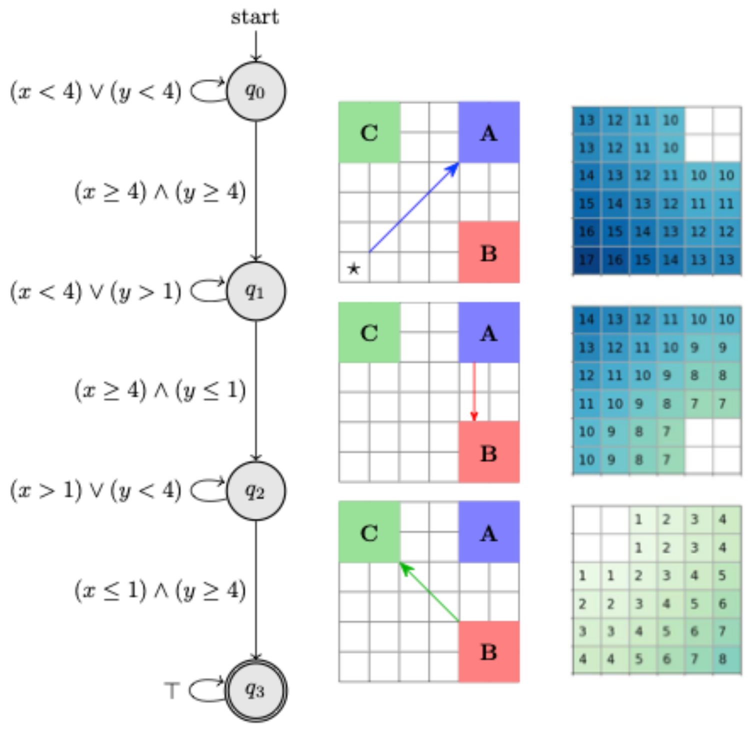
Open problems and proposed work:

Sim2real gap, data corruption, …

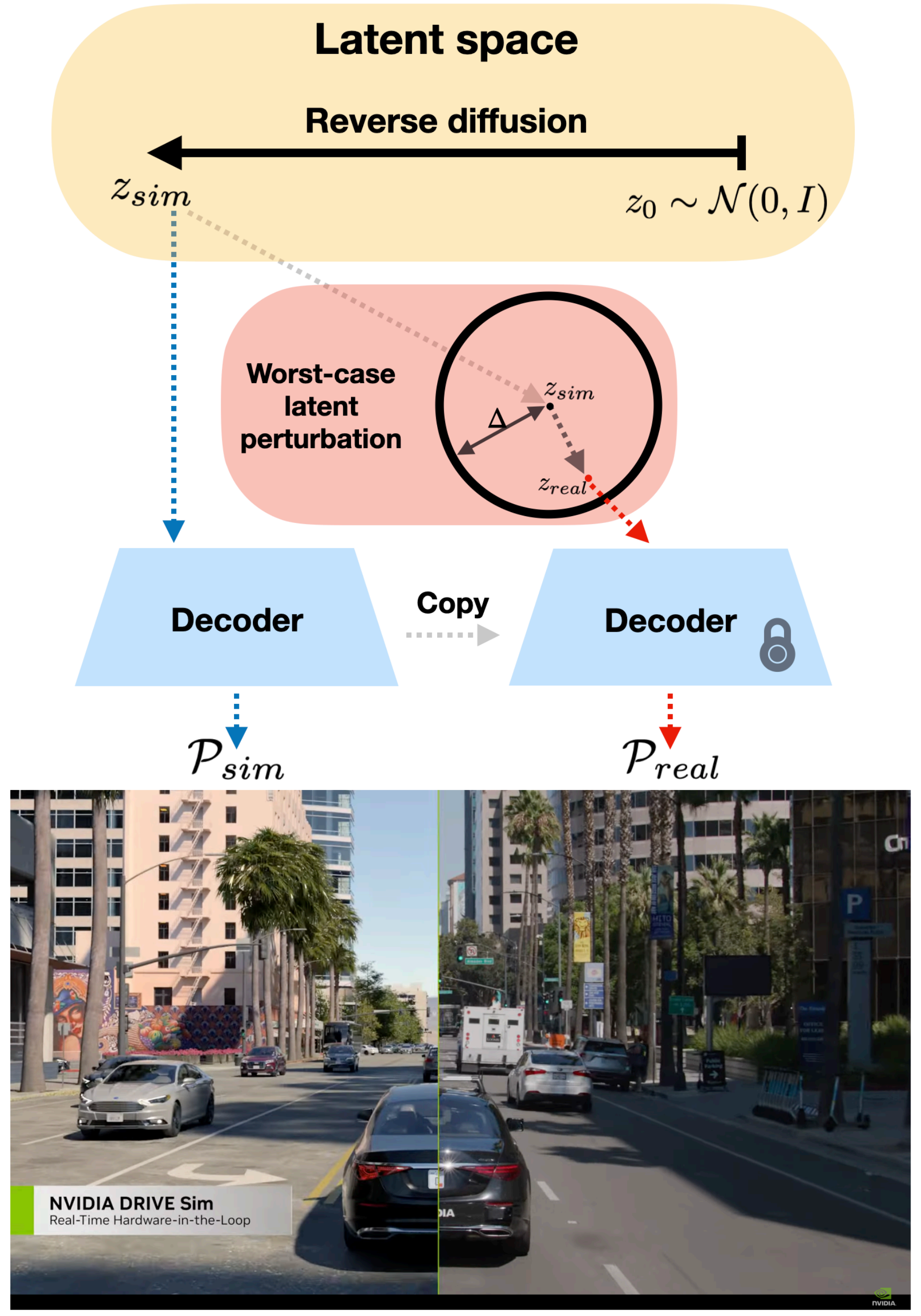
(3) Reinforcement learning for STL-GO

How to avoid reward hacking in LE-MASs?

- Decentralized MARL with STL-GO rewards
- Sample efficiency via symbolic automata



(4) Robust LEC training under distribution shift



- Learning algorithms robust to data corruption
- Distributionally robust learning via latent diffusion
- Robust predictive runtime monitoring and policy adaptation

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

(5) LE-MASs simulators for high quality data and real world validation

Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

Our statistical approach relies on realistic yet cheap data!

(5) LE-MASs simulators for high quality data and real world validation

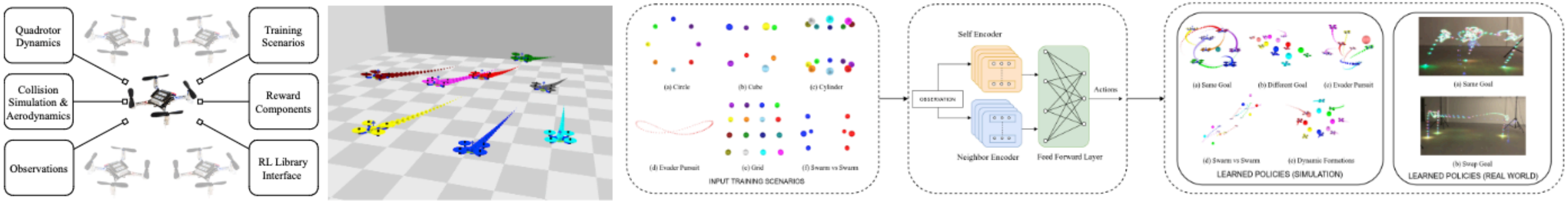
Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

Our statistical approach relies on realistic yet cheap data!

(5) LE-MASs simulators for high quality data and real world validation



QuadSwarm: physics-accurate, photo-realistic, parallelizable, and ready for integration

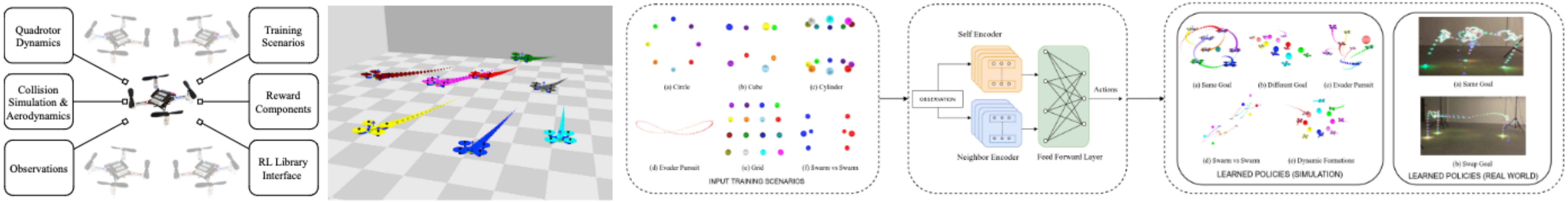
Statistical neurosymbolic reasoning for LE-MASs

Our approach is **neurosymbolic** as we combine rigorous symbolic reasoning with efficient neural network representations and **statistical** as we provide probabilistic end-to-end guarantees.

Open problems and proposed work:

Our statistical approach relies on realistic yet cheap data!

(5) LE-MASs simulators for high quality data and real world validation



QuadSwarm: physics-accurate, photo-realistic, parallelizable, and ready for integration

NVIDIA Drive and Isaac Sim: