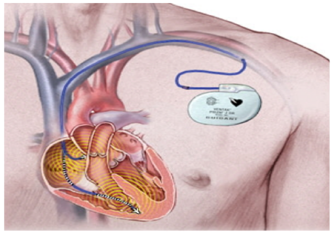
Context Aware Safety Monitoring in Medical Cyber-Physical Systems

Homa Alemzadeh
Dependable Systems and Analytics Group
Electrical and Computer Engineering
CPS Link Lab

University of Virginia

Medical Cyber-Physical Systems

Pacemakers



Insulin Pumps

Wearable Monitors

Patient Monitors

Infusion Pumps

Defibrillators

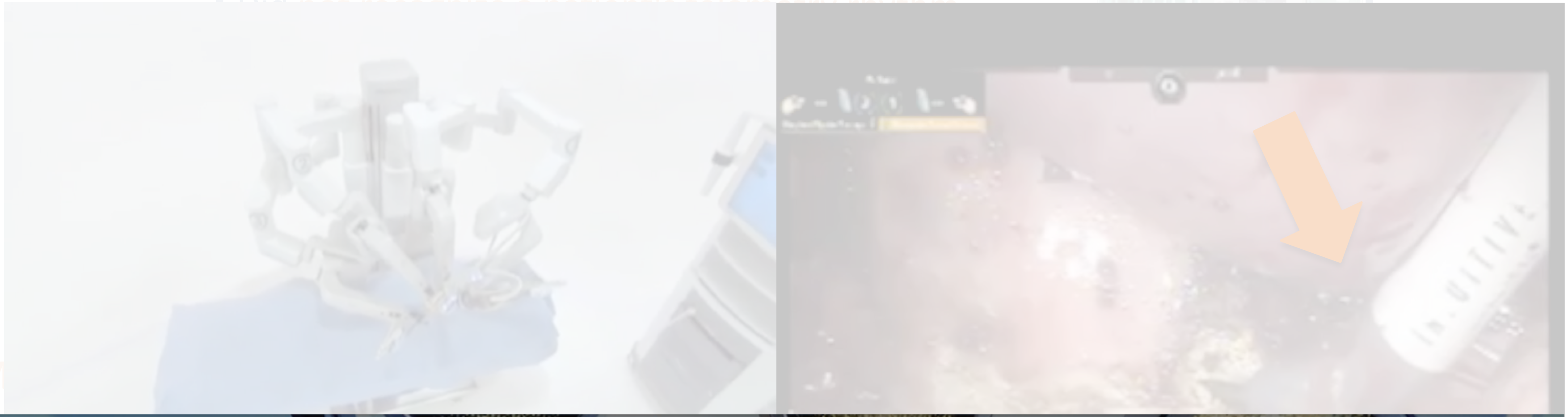
Surgical Robots

Imaging Systems

Linear Accelerators

Catastrophic Events

GE Healthcare - Telemetry Monitoring Systems

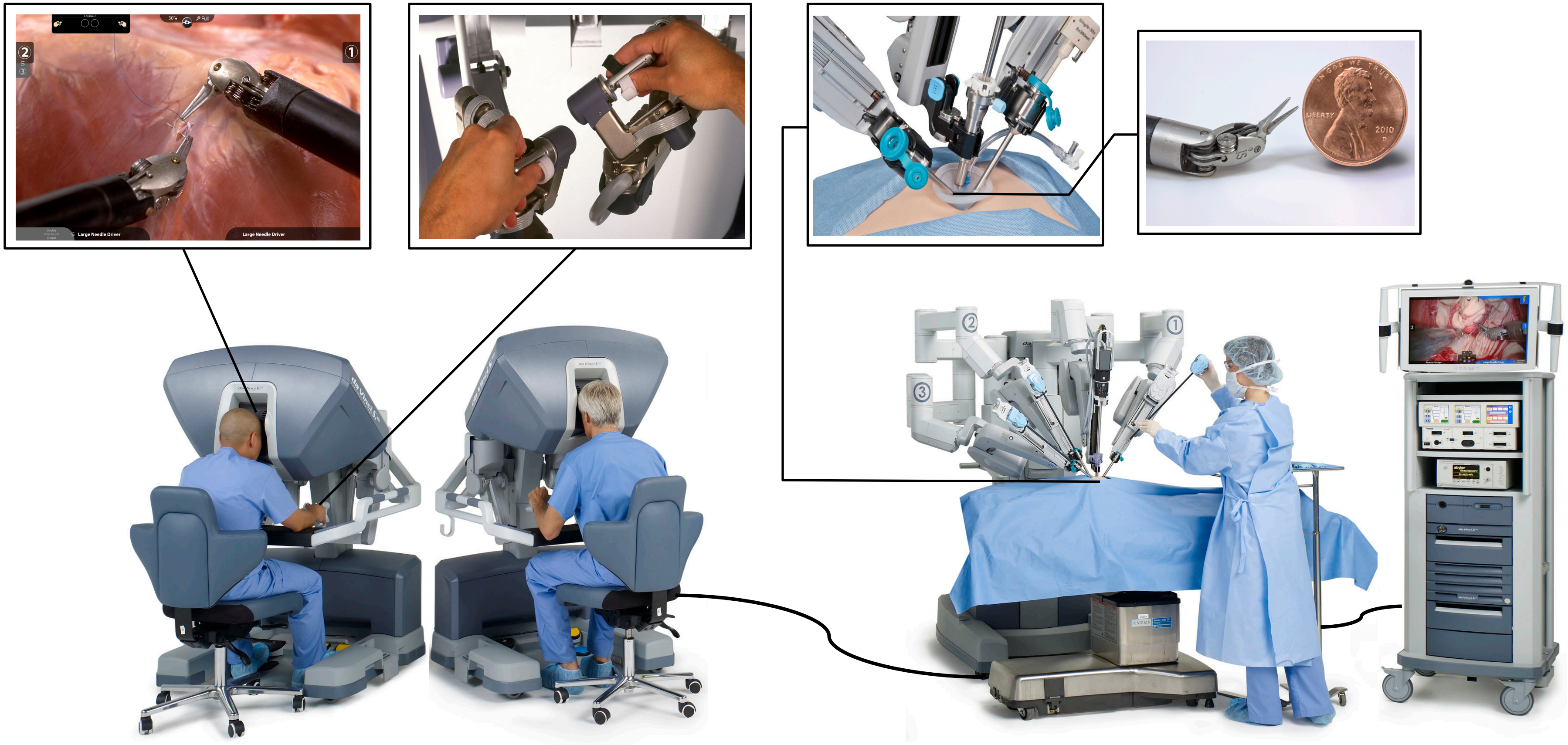


Hidden FDA Reports Detail Harm Caused By Scores Of Medical Devices

The Food and Drug Administration has let medical device companies file reports of injuries and malfunctions outside a widely scrutinized public database, which leave doctors and medical sleuths in the dark.

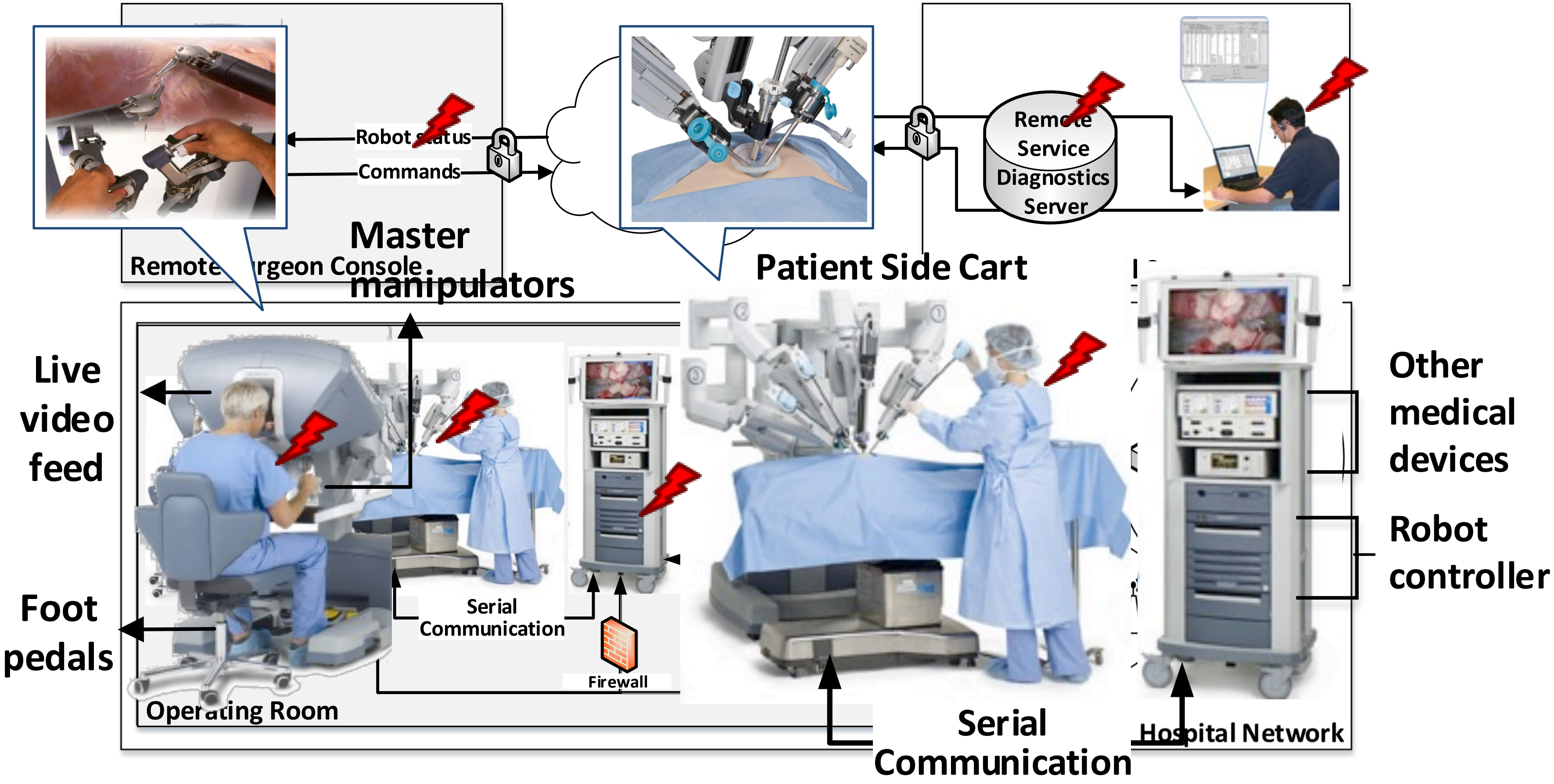
By **Christina Jewett** • Photos by **Heidi de Marco** • MARCH 7, 2019

Human-Cyber-Physical Systems



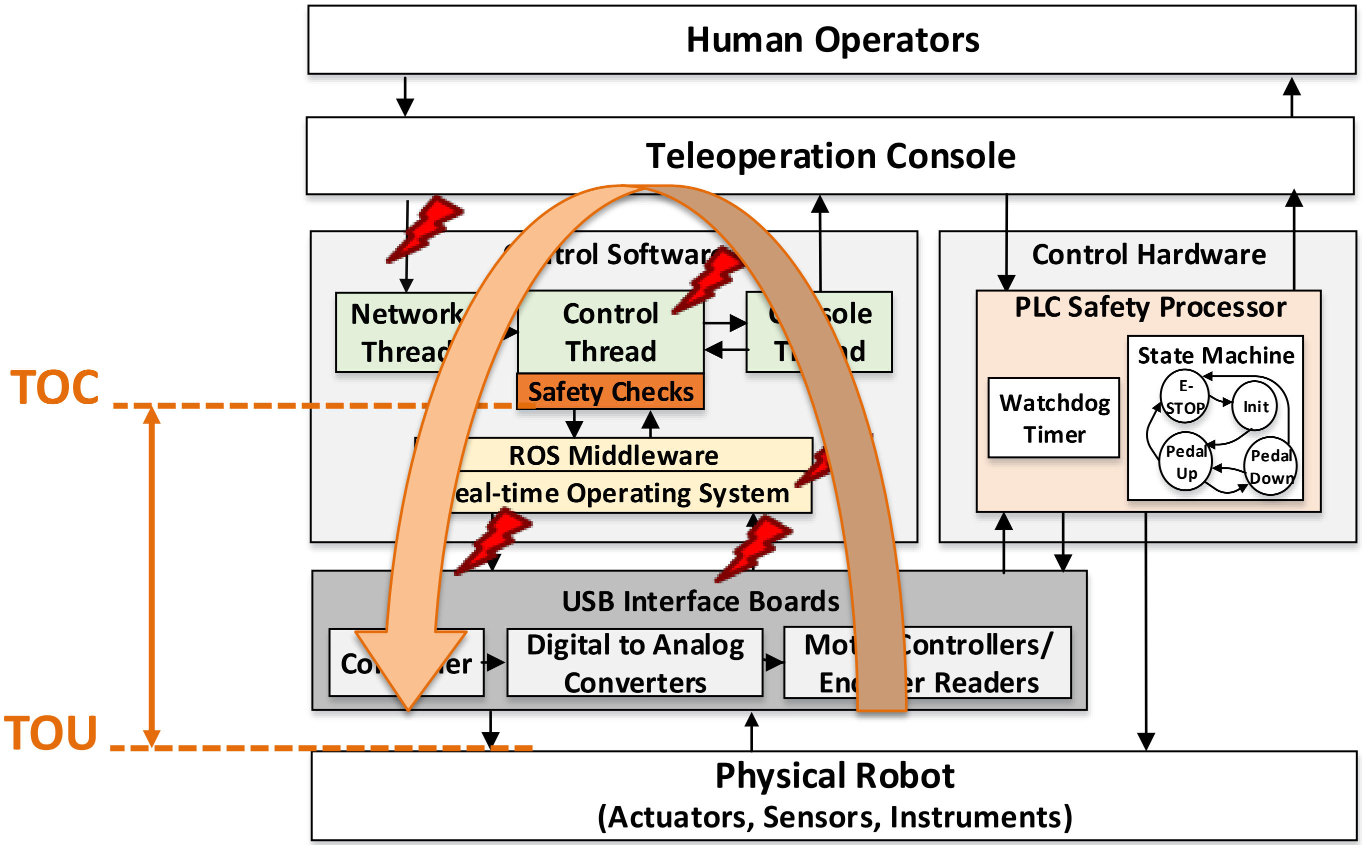
da Vinci Surgical System, Intuitive Surgical, Inc.

Human-Cyber-Physical Systems



da Vinci Surgical System, Intuitive Surgical, Inc.

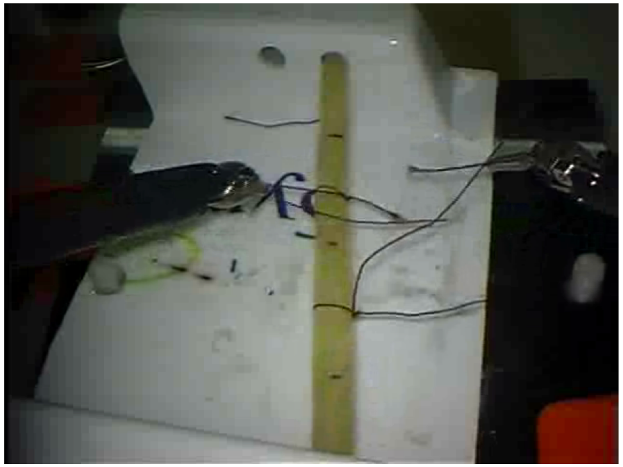
Vulnerabilities of Control System

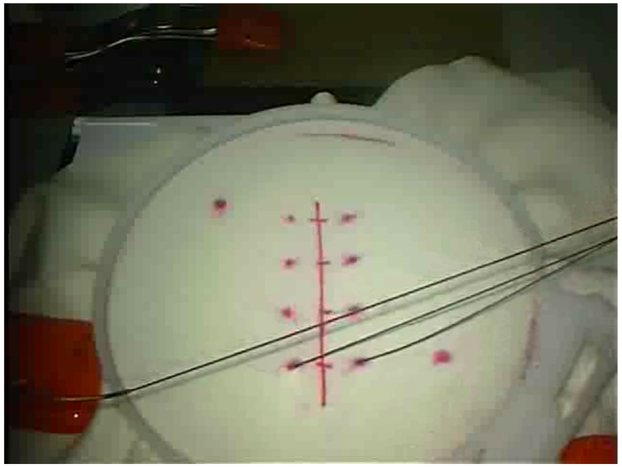


Each control cycle
1 milliseconds

RAVEN II, Applied Dexterity, University of Washington

Loosely Closed-loop System: No haptics, limited vision feedback



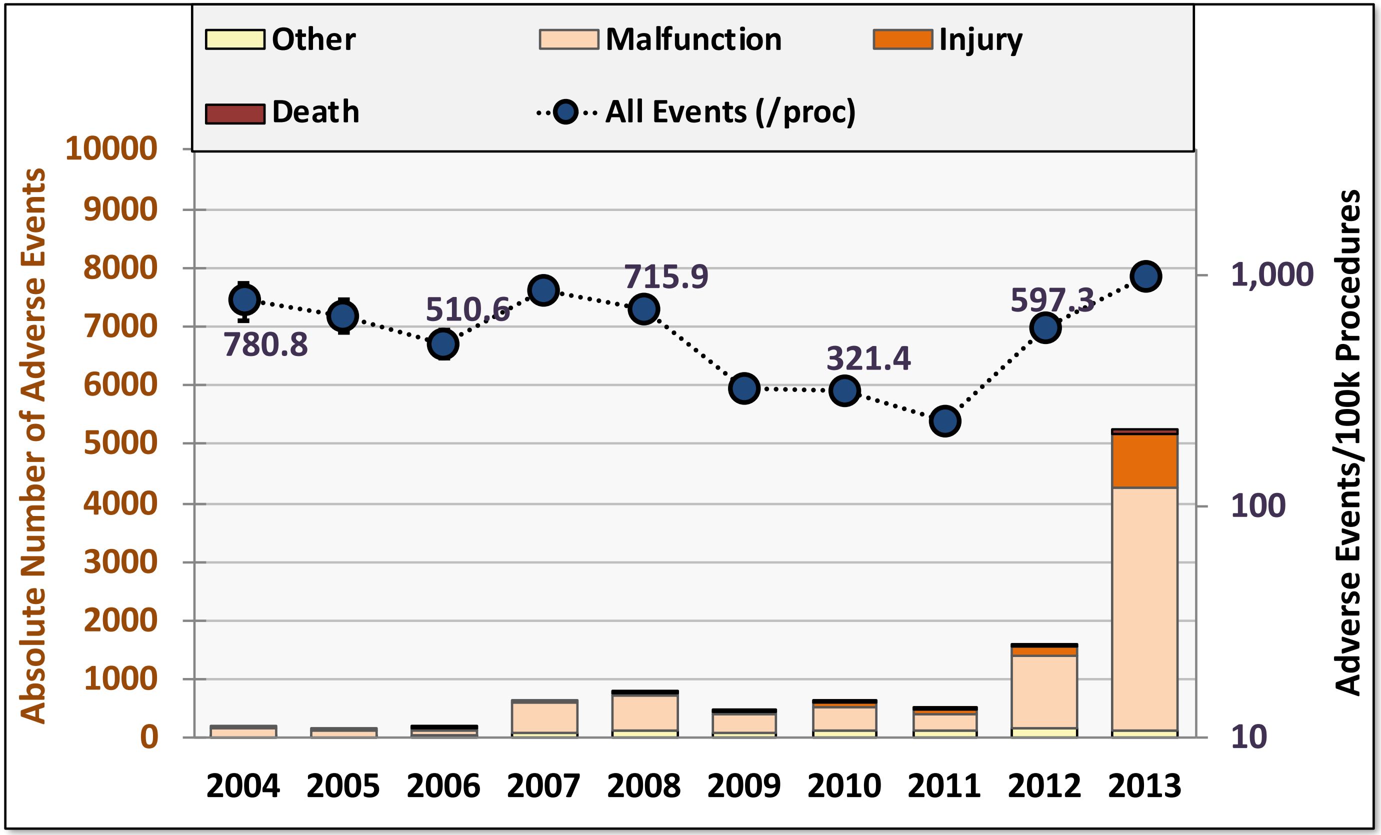


High Force

Graspers out of sight

Image Source: Gao, Yixin, et al., MICCAI Workshop: M2CAI. Vol. 3. 2014

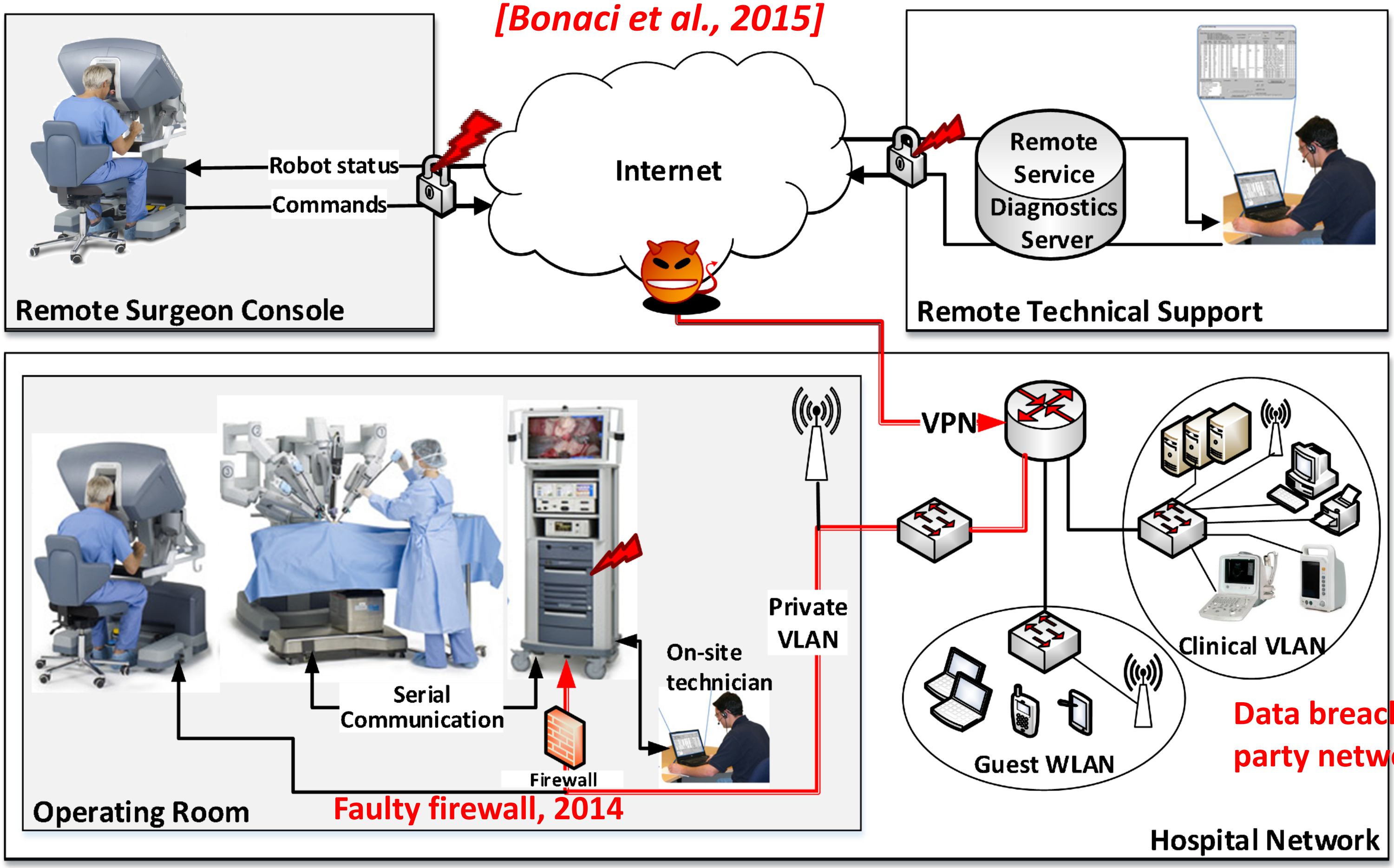
Once in every 100 procedures, an unexpected adverse event is likely to happen.



Alemzadeh et al., PLOS One 2016.
Featured in Wall Street Journal, MIT Technology Review, BBC, NBC News, Gizmodo, among others.

Malicious Attacks

DOS and MITM Attacks *[Bonaci et al., 2015]*



Faulty firewall, 2014

Email phishing, 2014

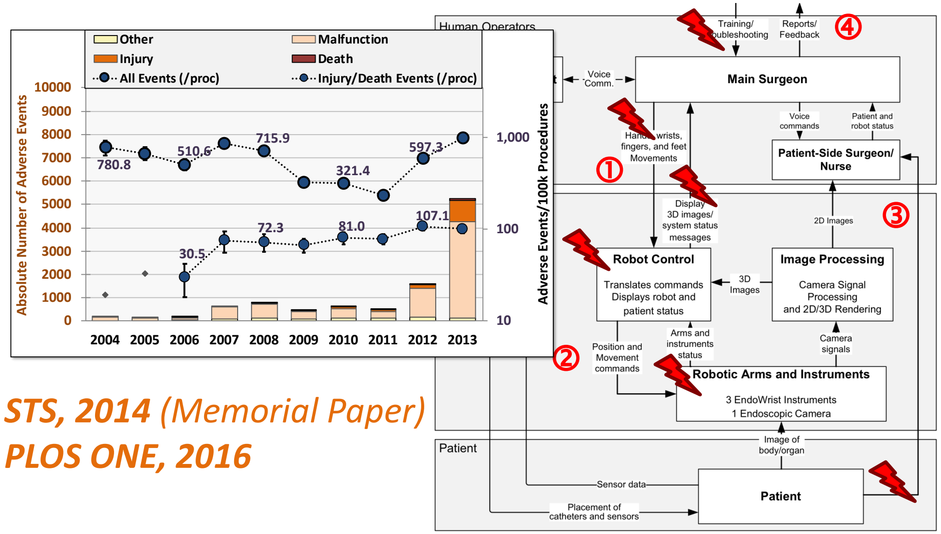
Default passwords, 2014

Medical devices, 2015

Data breach through 3rd party networks, 2015

Attacks on robot control system

Analysis of Adverse Events

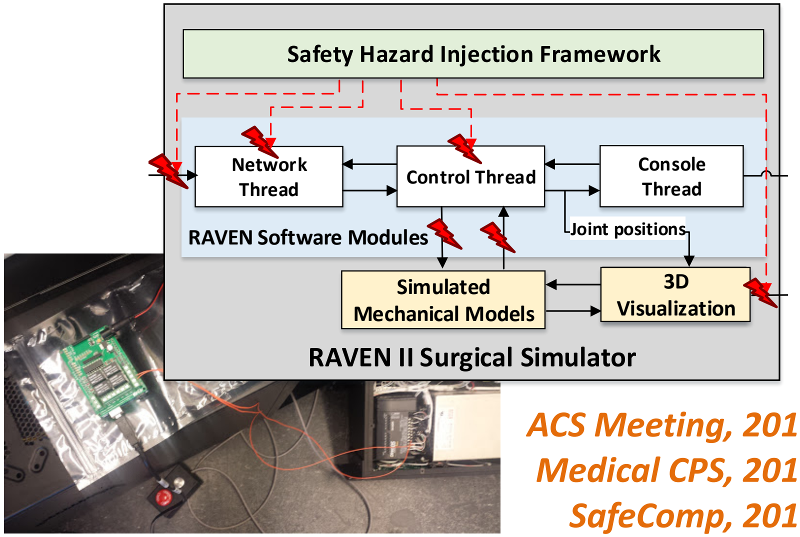


Human Operators
Training/ Troubleshooting
Reports/ Feedback
④
Voice Comm.
Main Surgeon
Voice commands
Patient and robot status
Hands, wrists, fingers, and feet Movements
①
Patient-Side Surgeon/ Nurse
Display 3D images/ system status messages
2D Images
③
Robot Control
Translates commands Displays robot and patient status
3D Images
Image Processing
Camera Signal Processing and 2D/3D Rendering
Position and Movement commands
Arms and instruments status
Camera signals
②
Robotic Arms and Instruments
3 EndoWrist Instruments
1 Endoscopic Camera
Patient
Image of body/organ
Sensor data
Patient
Placement of catheters and sensors

STS, 2014 (Memorial Paper)

PLOS ONE, 2016

System Resilience Assessment



Safety Hazard Injection Framework

Network Thread

Control Thread

Console Thread

RAVEN Software Modules

Joint positions

Simulated Mechanical Models

3D Visualization

RAVEN II Surgical Simulator

ACS Meeting, 2015

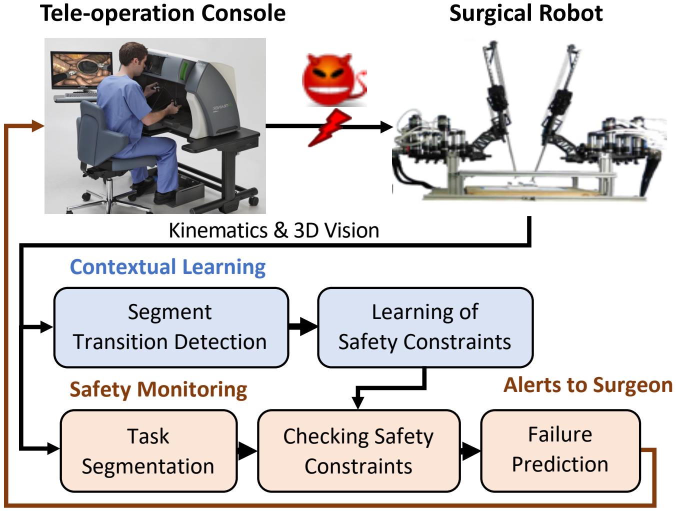
Medical CPS, 2015

SafeComp, 2015

IROS, 2016

PRDC, 2018

Real-time Safety Monitoring



Tele-operation Console

Surgical Robot

Kinematics & 3D Vision

Contextual Learning

Segment Transition Detection

Learning of Safety Constraints

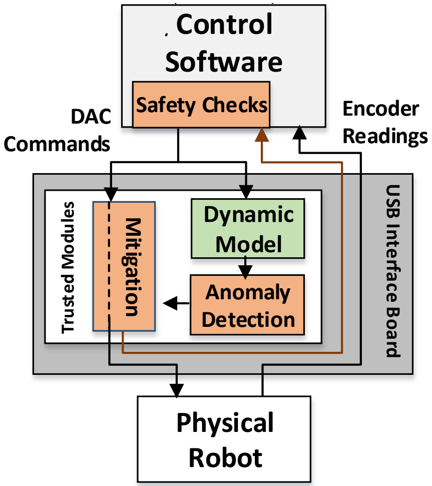
Safety Monitoring

Alerts to Surgeon

Task Segmentation

Checking Safety Constraints

Failure Prediction



Control Software

Safety Checks

DAC Commands

Encoder Readings

Trusted Modules

Mitigation

Dynamic Model

Anomaly Detection

USB Interface Board

Physical Robot

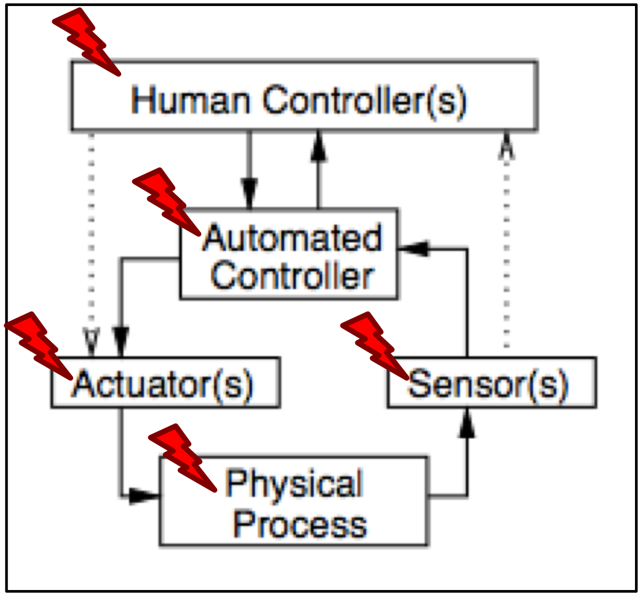
DSN, 2016

HotSoS, 2016

ISMR, 2019

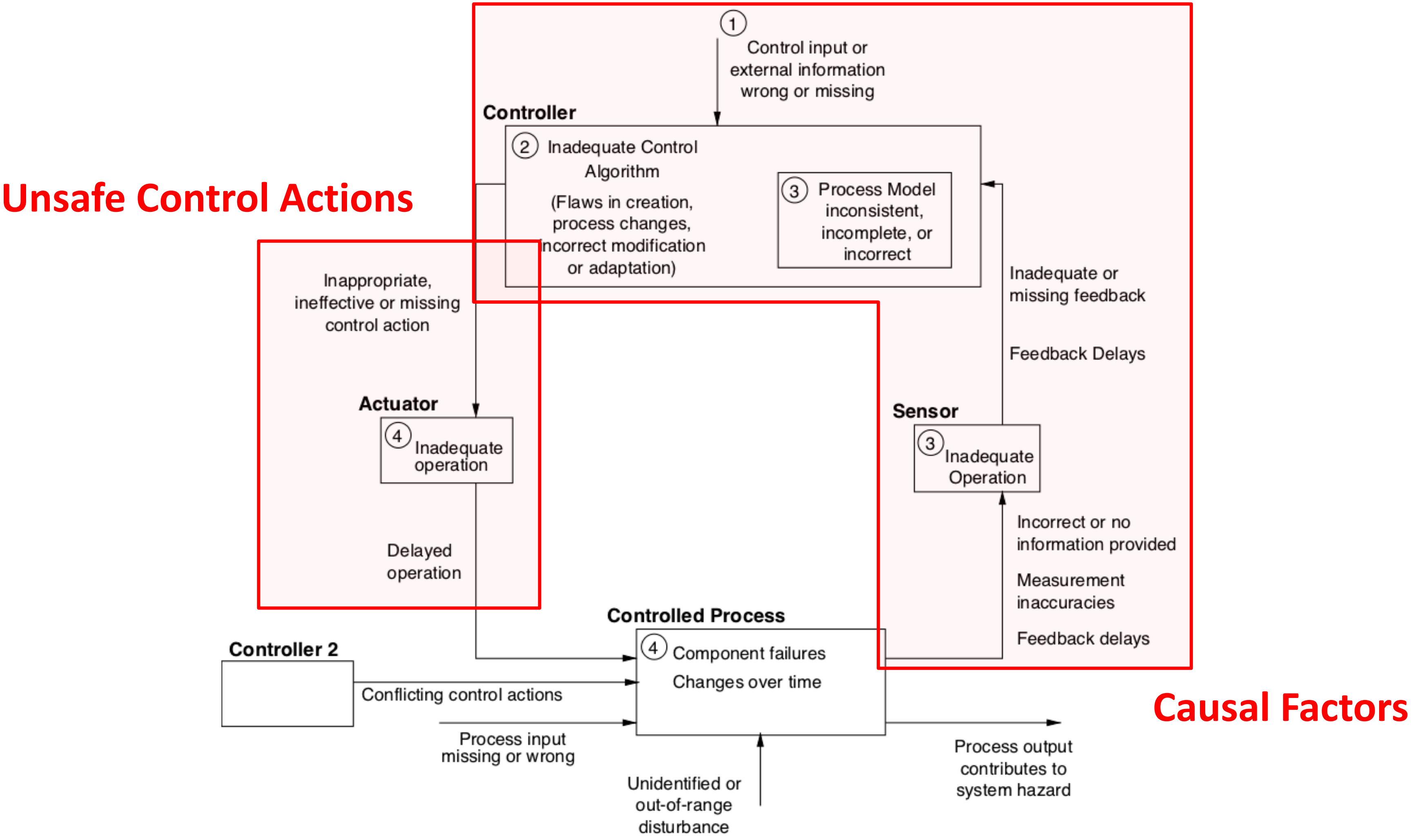
Real Time Safety Monitoring

- **Control-theoretic safety modeling and analysis**
  Violation of safety constraints in the control loops



Unexpected failures or intentional malicious actions leading to unsafe control

Unsafe Control Actions



Unsafe Control Actions

Causal Factors

Controller

1 Control input or external information wrong or missing

2 Inadequate Control Algorithm

(Flaws in creation, process changes, incorrect modification or adaptation)

3 Process Model inconsistent, incomplete, or incorrect

Inadequate or missing feedback

Feedback Delays

Inappropriate, ineffective or missing control action

Actuator

4 Inadequate operation

Delayed operation

Sensor

3 Inadequate Operation

Incorrect or no information provided

Measurement inaccuracies

Feedback delays

Controlled Process

4 Component failures

Changes over time

Controller 2

Conflicting control actions

Process input missing or wrong

Unidentified or out-of-range disturbance

Process output contributes to system hazard

Nancy Leveson, *Engineering a safer world: Systems thinking applied to safety*. MIT press, 2011.

Unsafe System Context

The set of system conditions under which the control actions could possibly be unsafe and lead to hazards.

i) a required control action was *not performed*

ii) a control action was performed *in a wrong state*

iii) a control action was performed *at an incorrect time*,

iv) a control action was performed *for an incorrect duration*,

v) a control action was provided, but *not followed by the controlled process*

Nancy Leveson, *Engineering a safer world: Systems thinking applied to safety*. MIT press, 2011.

Accidents and Safety Hazards

Accidents:

A-1. Patient expires during or after the procedure.

A-2. Patient is injured or experiences complications during/after the procedure.

A-3. Surgical system or instruments are damaged or lost.

Hazards:

H-1. Robot arms/instruments move:

- to unintended location (H1-1),
- with unintended velocity (H1-2),
- at unintended time (H1-3).

H-2. Robotic arms or instruments are subjected to collision/unintended stress.

H-3. Robotic system becomes unavailable or unresponsive during procedure.

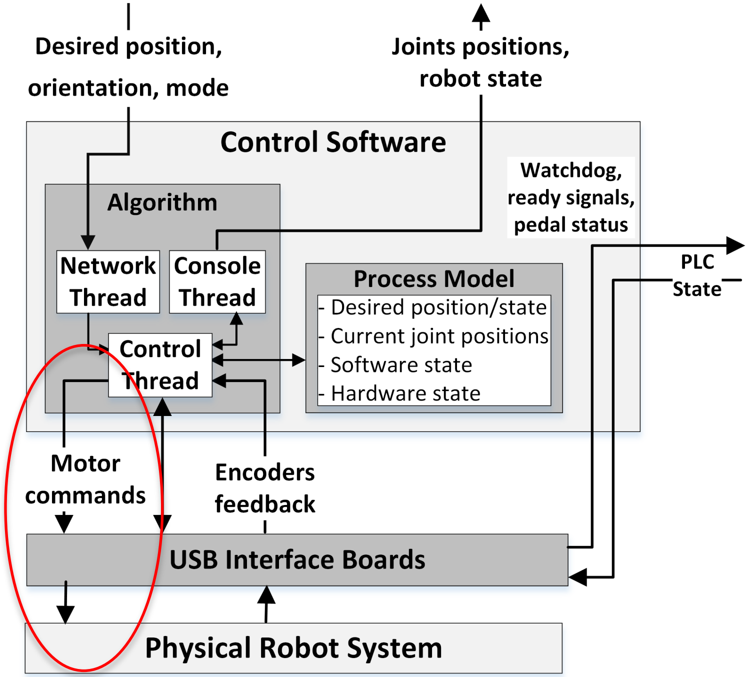
Unsafe System Context

ii) a control action was performed *in a wrong state*

A motor command is ***provided*** by control software when the ***user desired joint position is at a large distance from the current joint position***

***Potential hazard:* H1-2**

Robot arms/instruments will move with an unintended velocity



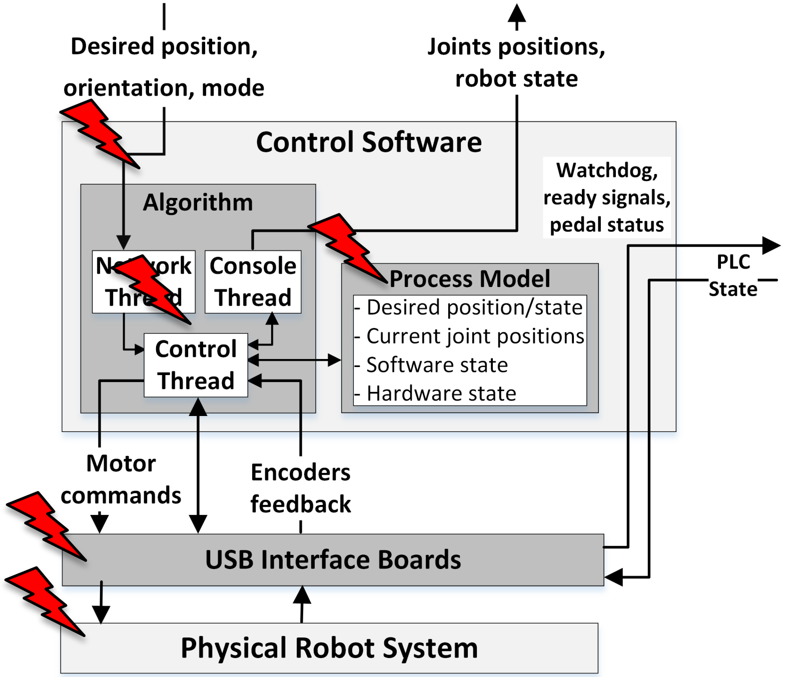
Unsafe Control Actions

ii) a control action was performed *in a wrong state*

A motor command is ***provided*** by control software when the ***user desired joint position is at a large distance from the current joint position***

Potential causes:

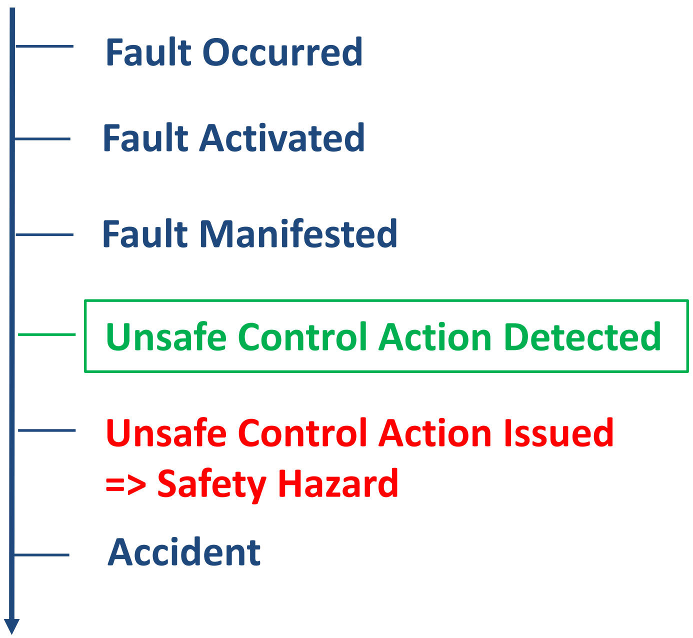
- Incorrect console inputs
- Faulty control algorithm
- Incorrect process model
- Faulty USB communication
- Physical system malfunction



Real Time Safety Monitoring

- **Preemptive Detection of Safety Hazards**
  Unsafe system context leading to unsafe control actions

Fault Propagation Timeline



Fault Occurred

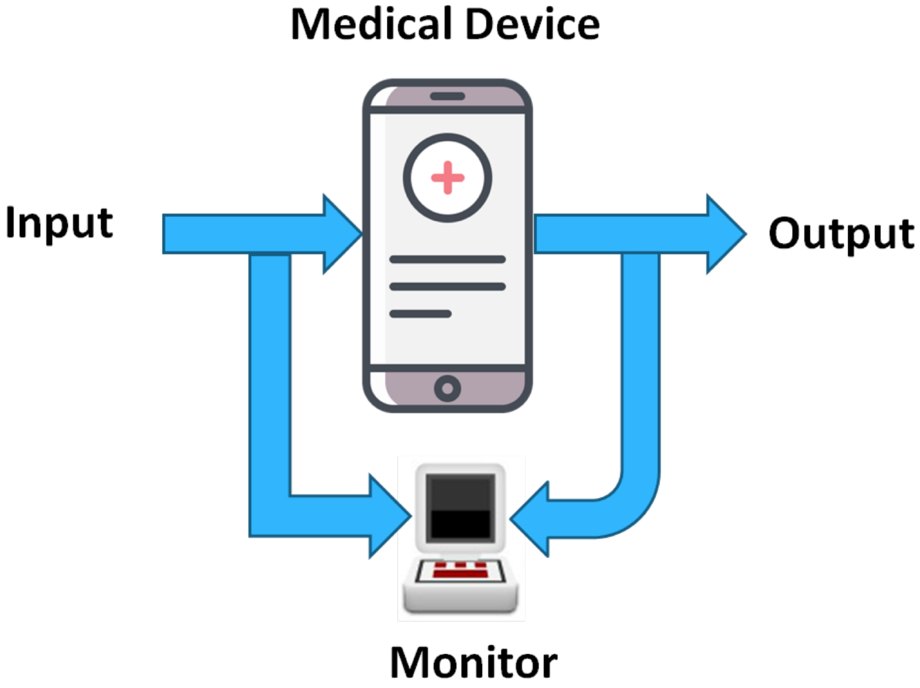
Fault Activated

Fault Manifested

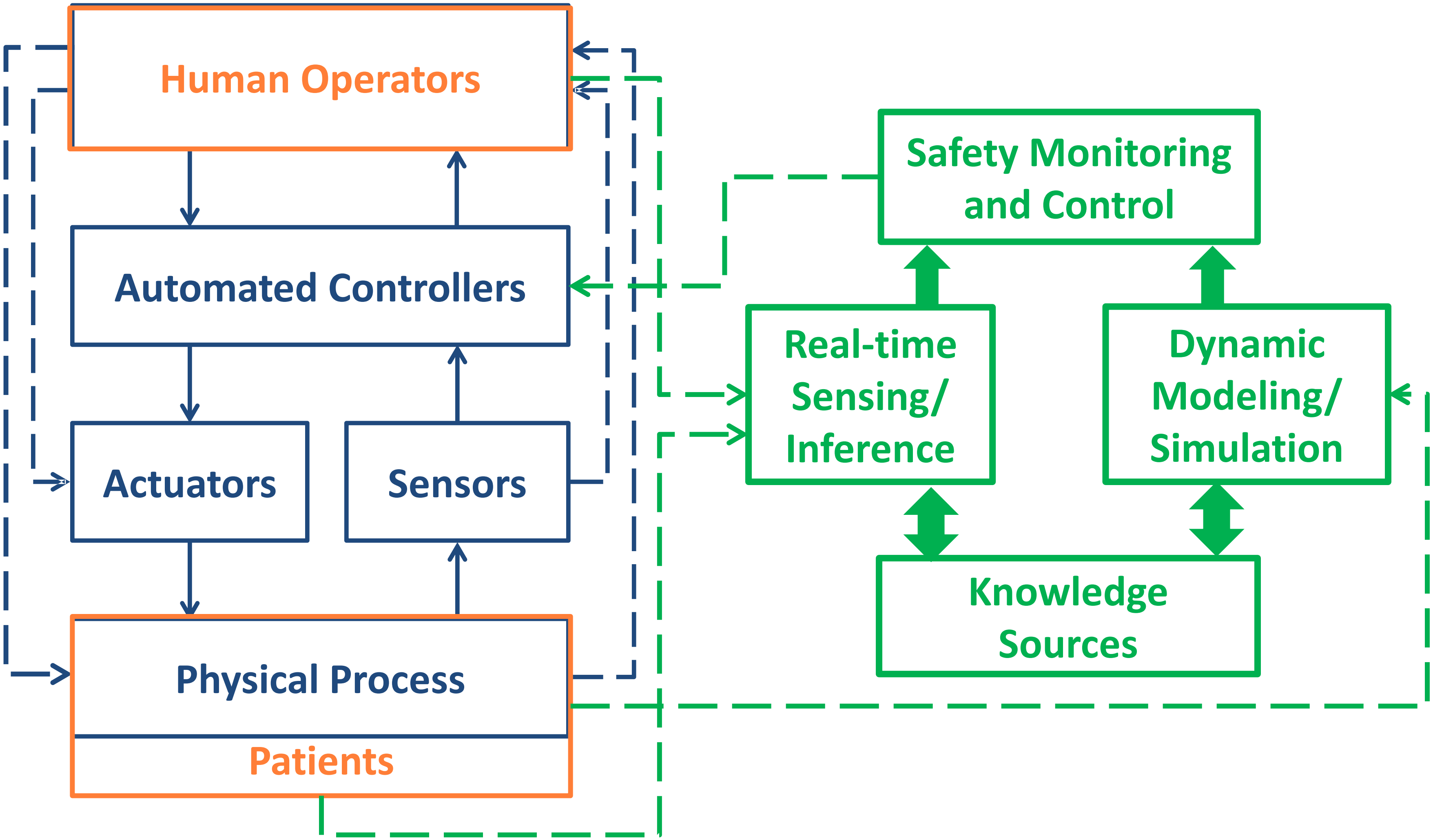
Unsafe Control Action Detected

Unsafe Control Action Issued
=> Safety Hazard

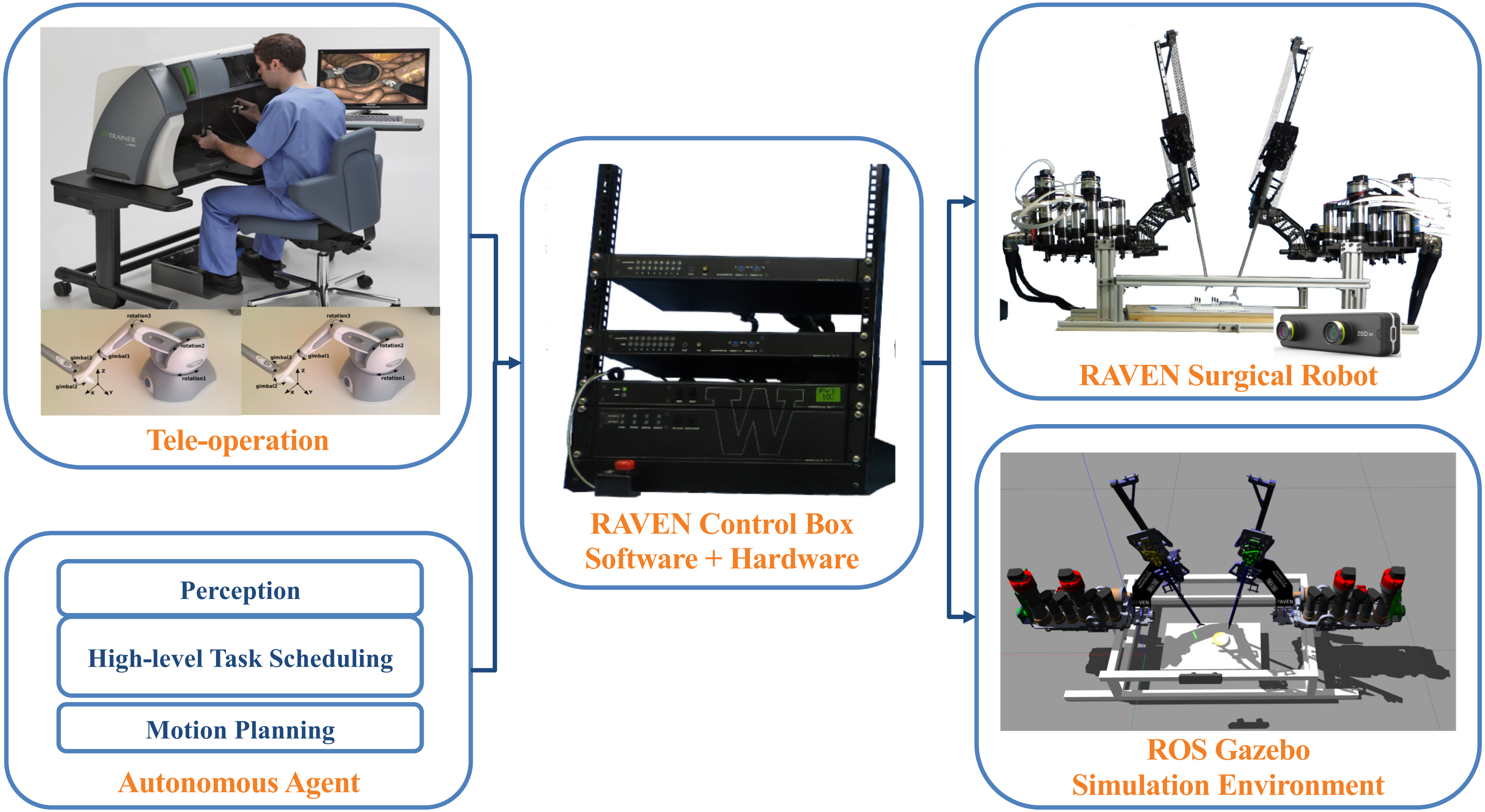
Accident



Context-Aware Safety Monitoring



RAVEN II Surgical Robot



Tele-operation

Perception

High-level Task Scheduling

Motion Planning

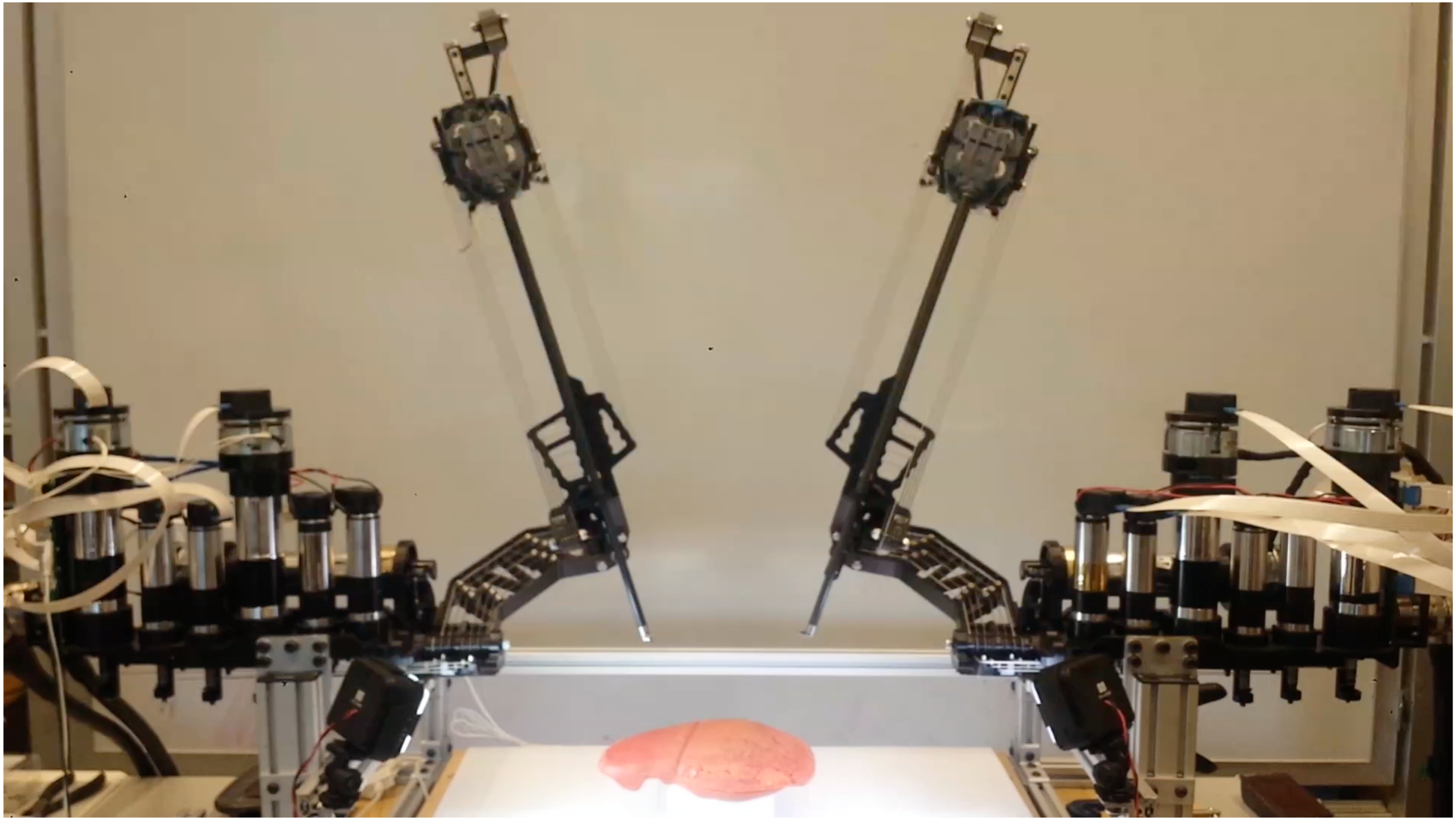
Autonomous Agent

RAVEN Control Box
Software + Hardware

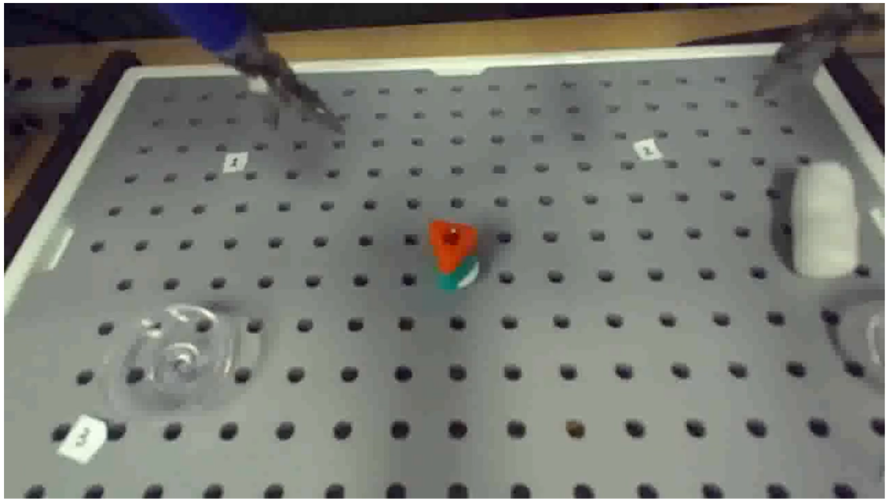
RAVEN Surgical Robot

ROS Gazebo
Simulation Environment

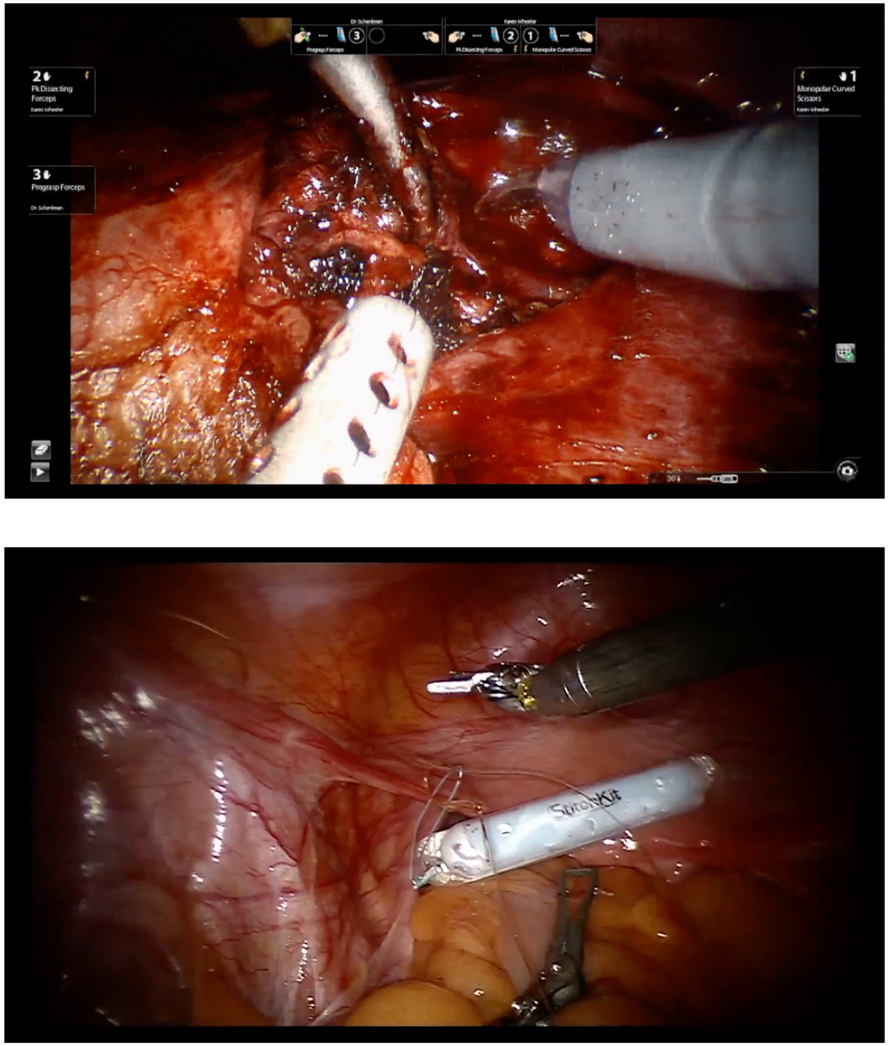
RAVEN II Surgical Robot



Example Surgical Task: Pick and Place

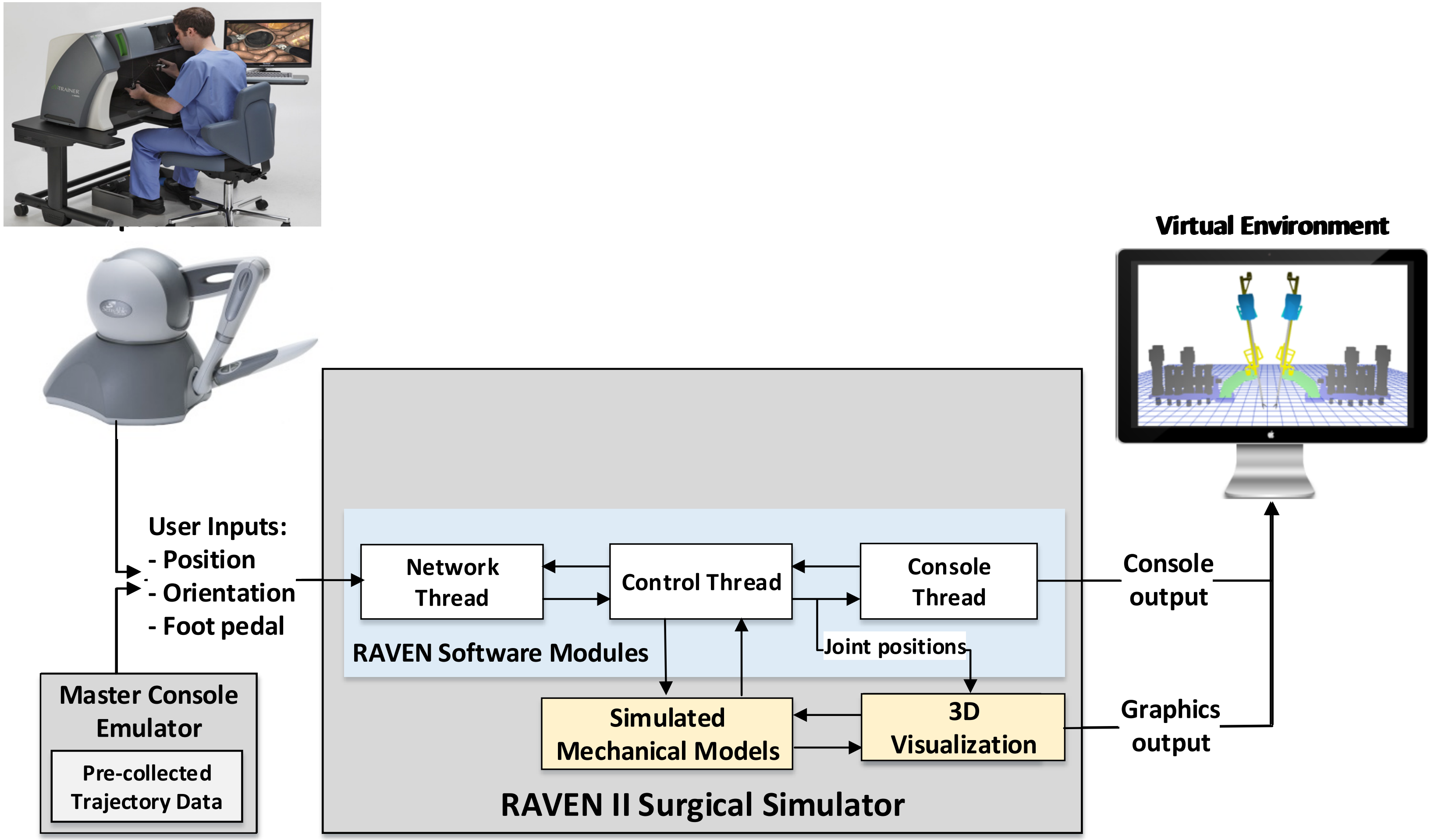


Dry Lab Simulation



Actual Surgery

Simulator + Hazard Injection Engine



Virtual Environment

User Inputs:
- Position
- Orientation
- Foot pedal

Master Console Emulator

Pre-collected Trajectory Data

Network Thread

Control Thread

Console Thread

Joint positions

RAVEN Software Modules

Simulated Mechanical Models

3D Visualization

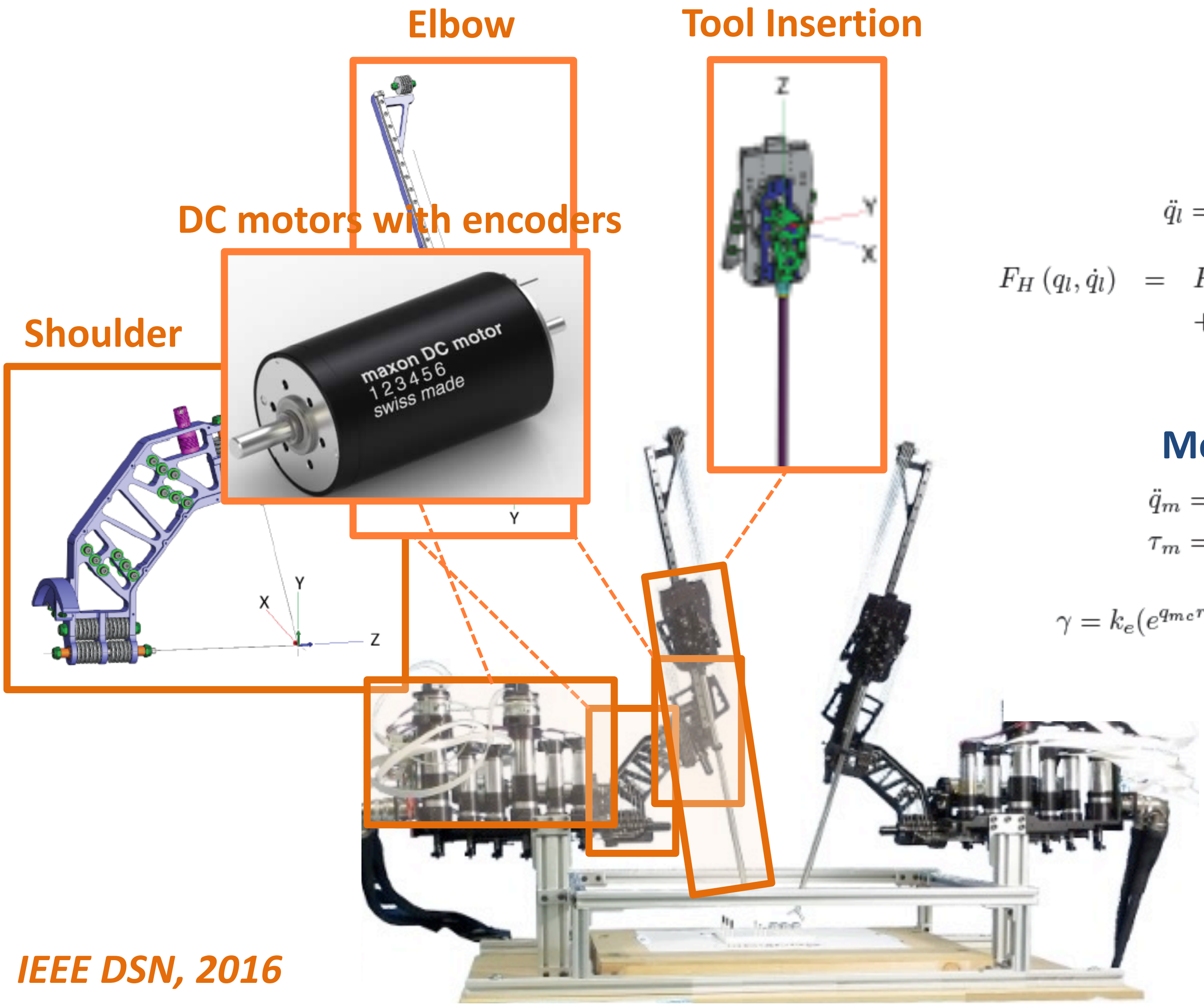
RAVEN II Surgical Simulator

Console output

Graphics output

https://github.com/UVA-DSA/raven2_sim

Joint and Motor Dynamics



Joint Dynamics:

$$\ddot{q}_l = {I_l}^{-1}\left[\Gamma - F_H\left(q_l, \dot{q}_l\right)\right] \tag{1}$$

$$\begin{aligned} F_H\left(q_l, \dot{q}_l\right) \quad = \quad & F_C\left(q_l, \dot{q}_l\right) + F_G + \mathrm{diag}\left(sign(\dot{q}_l)\right) F_{cl} \\ & + \mathrm{diag}\left(\dot{q}_l\right) F_{vl} + J^T F_{en} \end{aligned} \tag{2}$$

Motor and Cable Dynamics:

$$\ddot{q}_m = (1/I_m)(\tau - \tau_m - \tau_{rn}) \tag{3}$$

$$\tau_m = \tau_{cm}\mathrm{sign}(\dot{q_m}) + \tau_{vm}\dot{q_m} \tag{4}$$

$$\tau_{rn} = r_{mc}\gamma/N \tag{5}$$

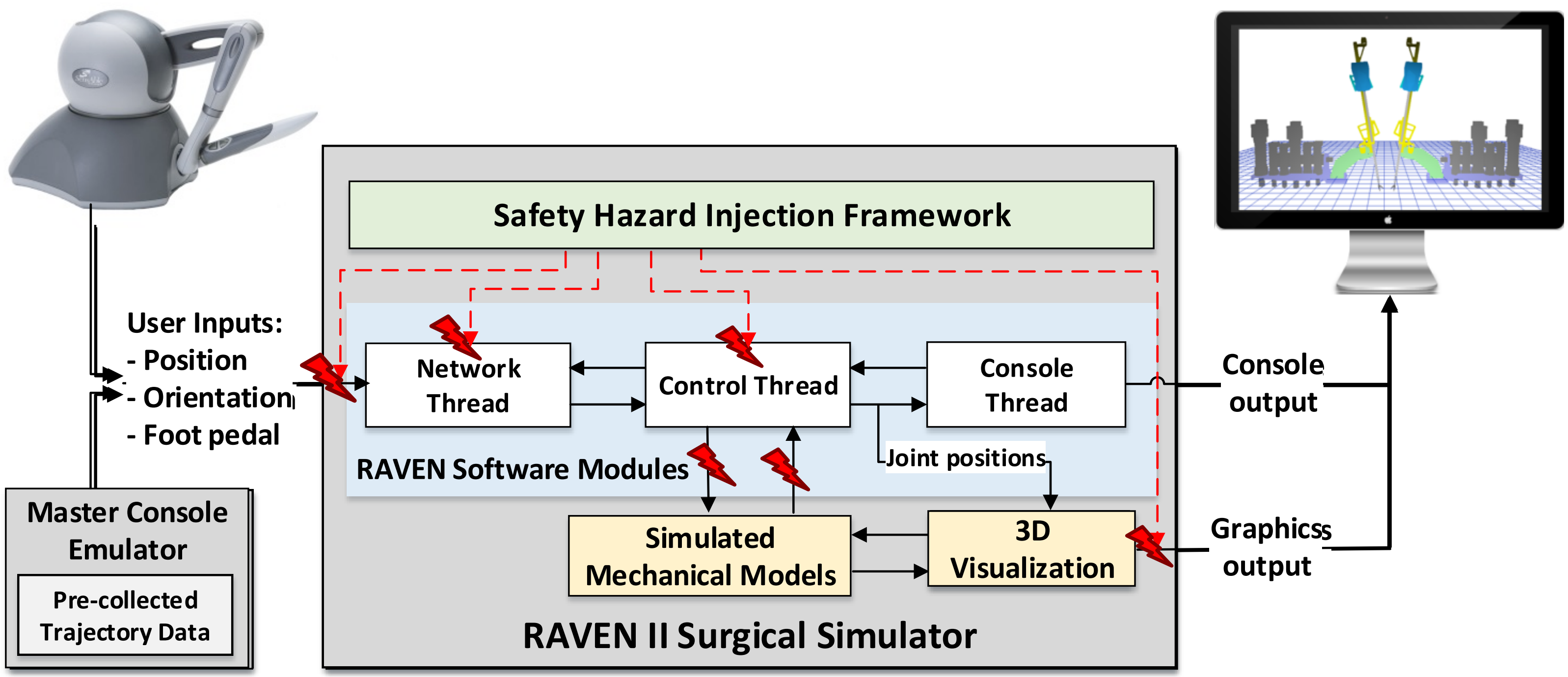
$$\begin{aligned} \gamma = k_e(e^{q_{mc}r_{mc} - q_l r_l} - e^{q_l r_l - q_{mc}r_{mc}}) \\ + 2b_e(\dot{q_{mc}}r_{mc} - \dot{q}_l r_l) \end{aligned} \tag{6}$$

$$\Gamma = r_l\gamma \tag{7}$$

IEEE DSN, 2016

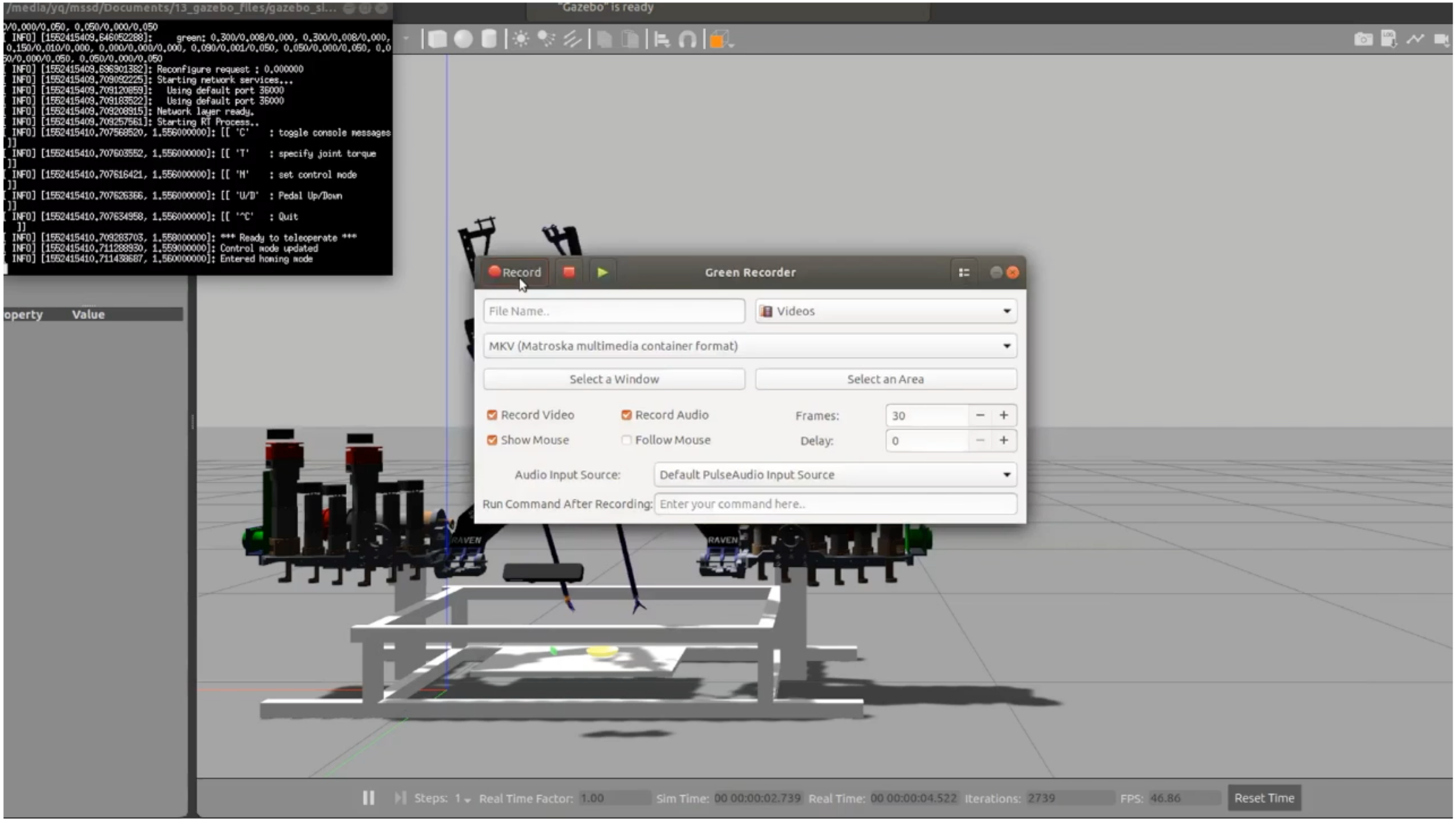
Haghighipanah, et al., 2015

Simulator + Hazard Injection Engine



https://github.com/UVA-DSA/raven2_sim

Pick and Place Task in Gazebo Simulator



/media/yq/mssd/Documents/13_gazebo_files/gazebo_sl...

```
0/0.000/0.050, 0.050/0.000/0.050
[ INFO] [1552415409.646052288]:     green: 0.300/0.008/0.000, 0.300/0.008/0.000,
 0.150/0.010/0.000, 0.000/0.000/0.000, 0.090/0.001/0.050, 0.050/0.000/0.050, 0.0
50/0.000/0.050, 0.050/0.000/0.050
[ INFO] [1552415409.696901382]: Reconfigure request : 0.000000
[ INFO] [1552415409.709092225]: Starting network services...
[ INFO] [1552415409.709120859]:   Using default port 36000
[ INFO] [1552415409.709183522]:   Using default port 36000
[ INFO] [1552415409.709208915]: Network layer ready.
[ INFO] [1552415409.709257561]: Starting RT Process..
[ INFO] [1552415410.707568520, 1.556000000]: [[ 'C'    : toggle console messages
 ]]
[ INFO] [1552415410.707603552, 1.556000000]: [[ 'T'    : specify joint torque
 ]]
[ INFO] [1552415410.707616421, 1.556000000]: [[ 'M'    : set control mode
 ]]
[ INFO] [1552415410.707626366, 1.556000000]: [[ 'U/D'  : Pedal Up/Down
 ]]
[ INFO] [1552415410.707634958, 1.556000000]: [[ '^C'   : Quit
  ]]
[ INFO] [1552415410.709283703, 1.558000000]: *** Ready to teleoperate ***
[ INFO] [1552415410.711288930, 1.559000000]: Control mode updated
[ INFO] [1552415410.711438687, 1.560000000]: Entered homing mode
```

"Gazebo" is ready

Green Recorder

File Name..:
Videos
MKV (Matroska multimedia container format)
Select a Window | Select an Area

☑ Record Video ☑ Record Audio Frames: 30
☑ Show Mouse □ Follow Mouse Delay: 0

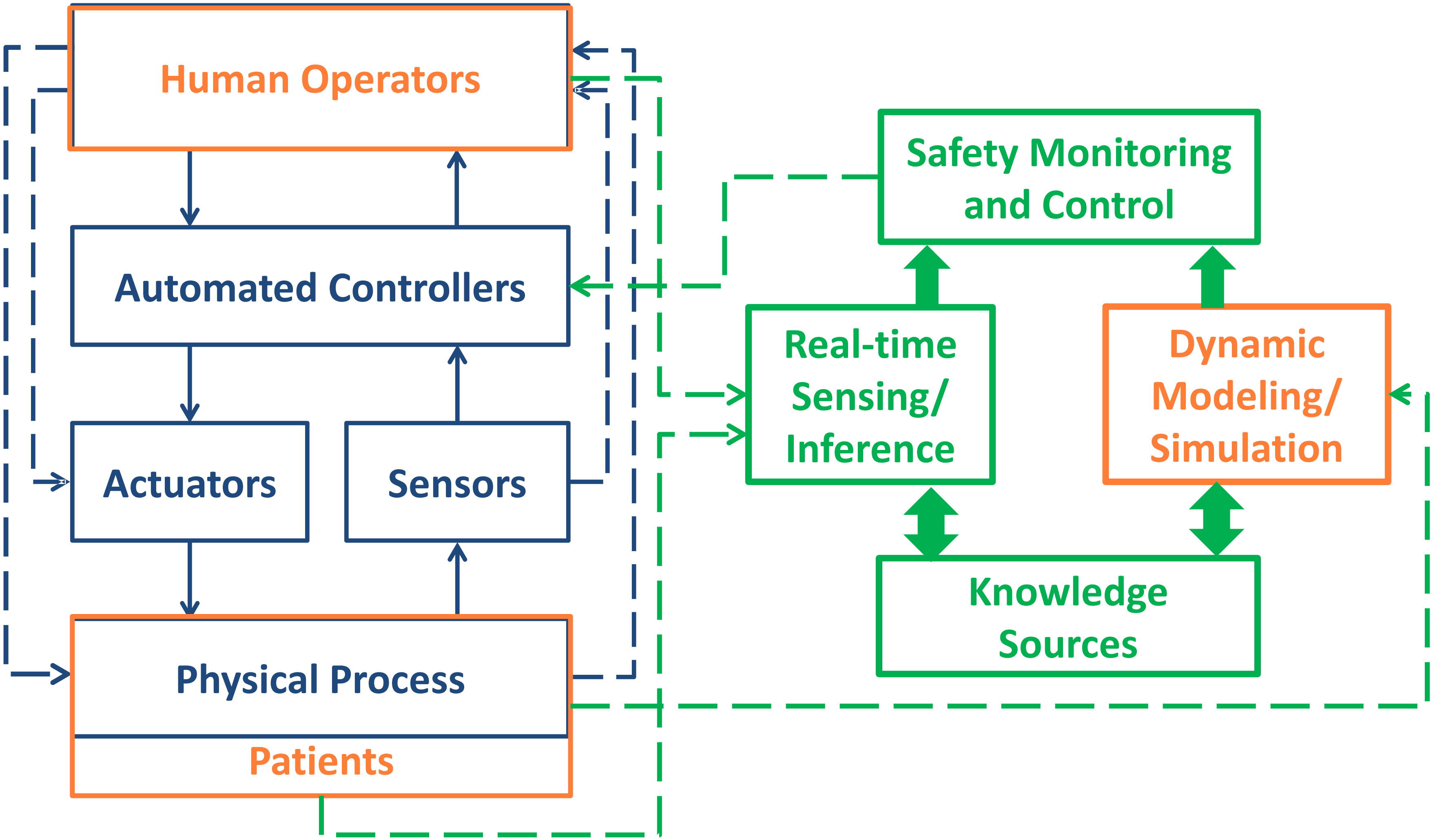
Audio Input Source: Default PulseAudio Input Source

Run Command After Recording: Enter your command here..

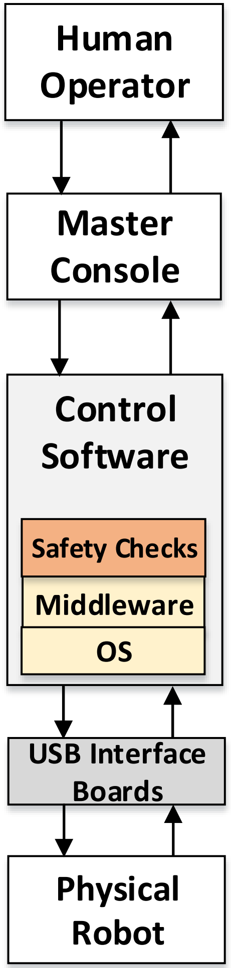
Steps: 1 Real Time Factor: 1.00 Sim Time: 00 00:00:02.739 Real Time: 00 00:00:04.522 Iterations: 2739 FPS: 46.86

Failure Modes in Gazebo Simulator

Context-Aware Safety Monitoring

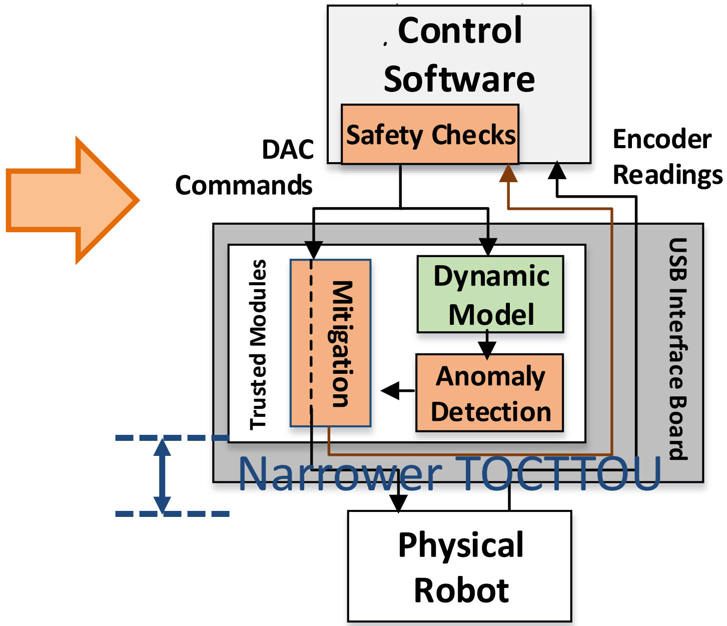


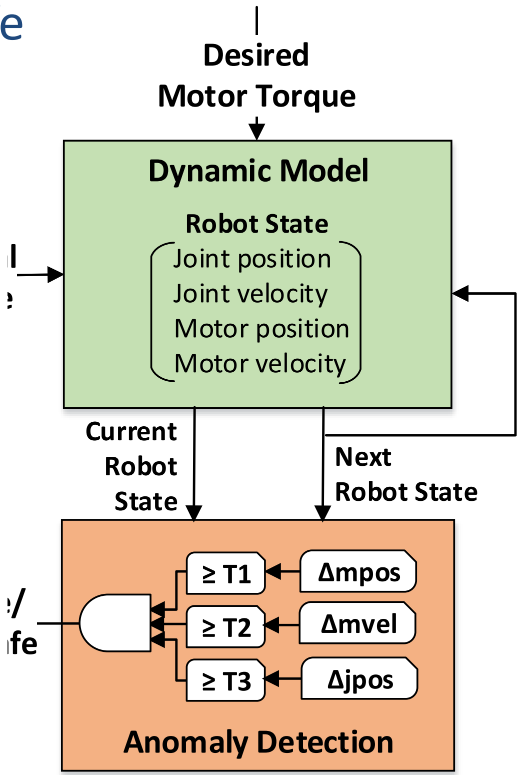
Dynamic Model Based Detection



Preemptive detection of safety hazards

- Real-time computation of joint/motor dynamics
- Estimation of next robot state
- Detect if distance is unsafe





< 1 ms

Safety Hazard Detection Performance

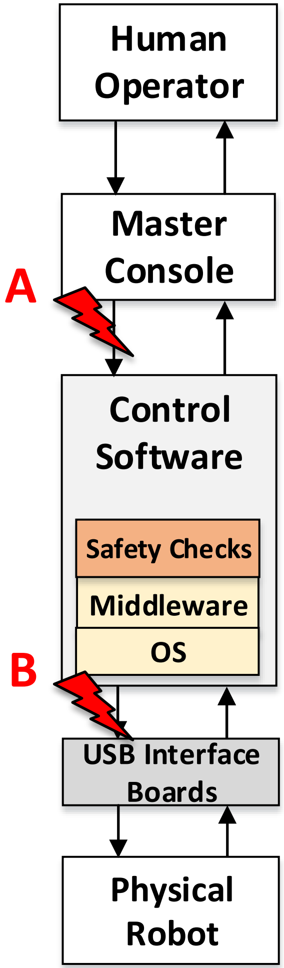
Simulated attack scenarios

Scenario A: 1,925 runs

Scenario B: 1,361 runs

Different injected error values and attack activation periods

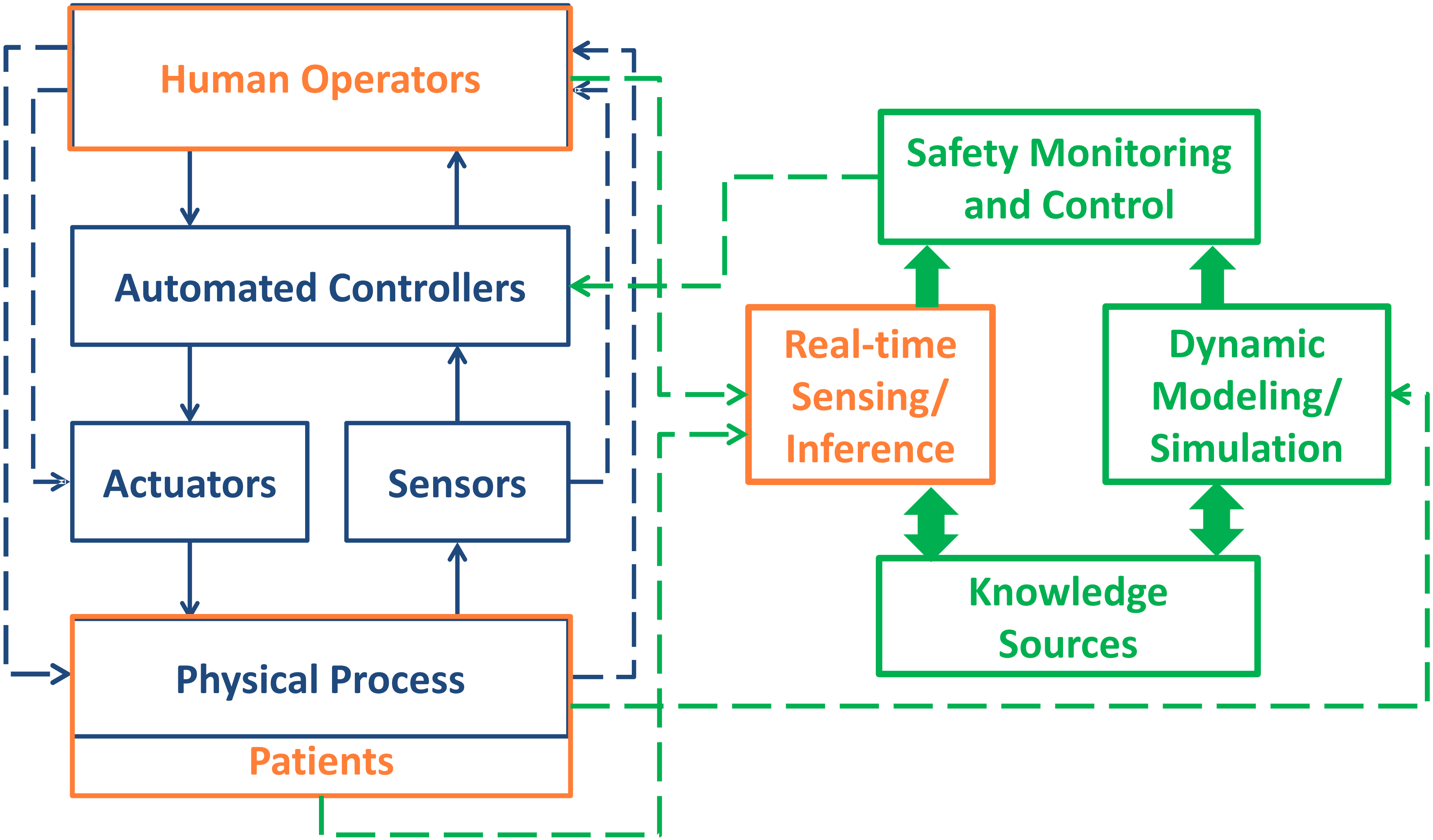
Attack Scenario	Technique	ACC (%)	TPR (%)	FPR (%)	F1 (%)
A (User inputs)	DM	88.0	89.8	12.4	74.8
	RAVEN	84.6	53.3	7.7	57.8
B (Torque commands)	DM	92.0	99.8	11.8	89.1
	RAVEN	90.7	81.0	4.6	85.1



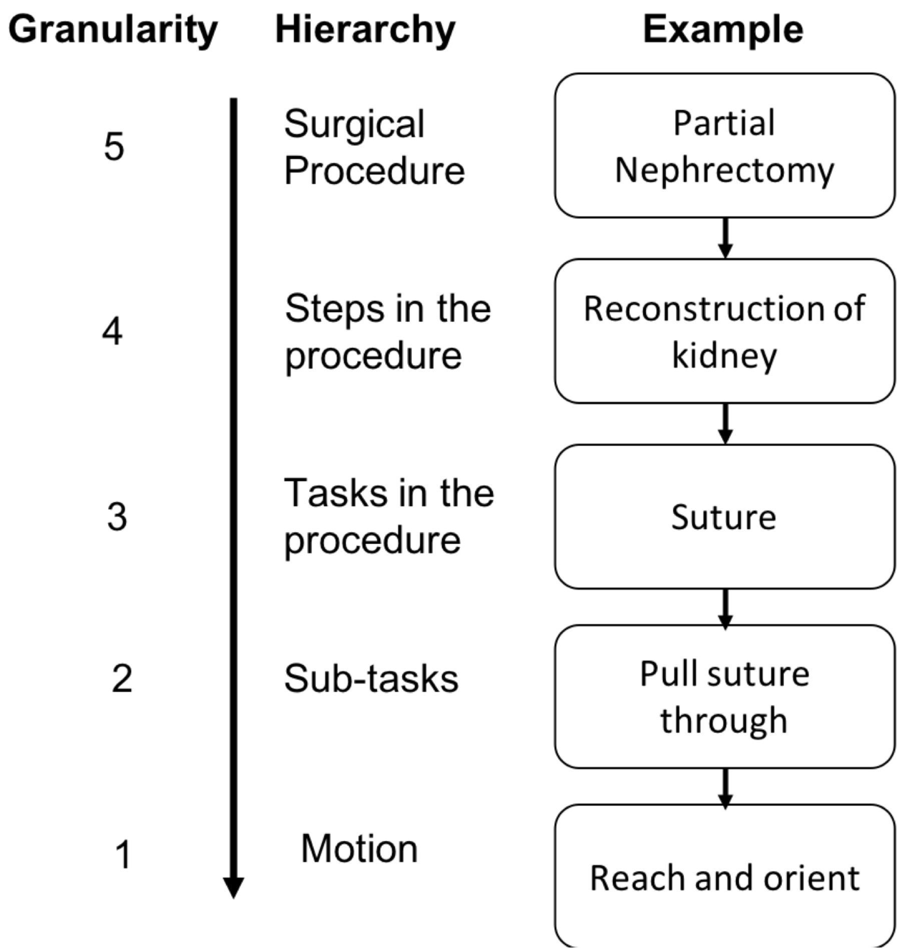
- DM detected **before** hazard manifested in physical layer
- RAVEN detected *at least* 1 cycle **after** safety hazard occurred

Alemzadeh et al., DSN 2016.

Context-Aware Safety Monitoring

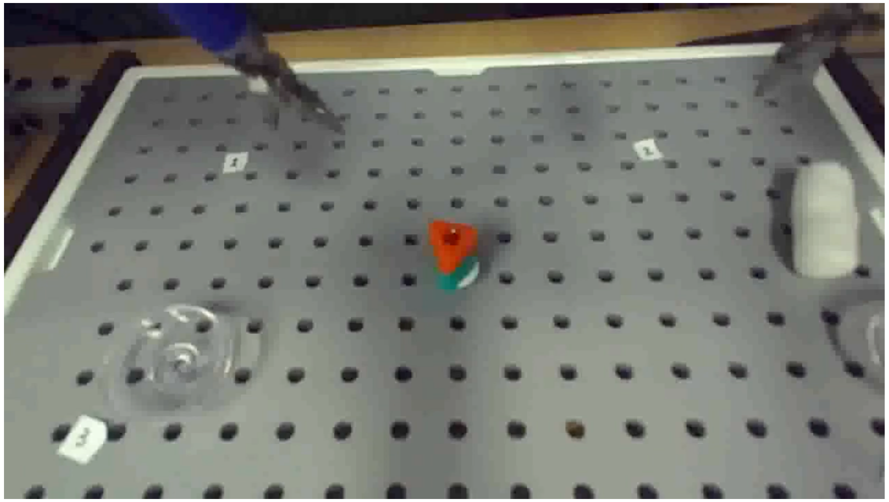


Operational Context in Surgery

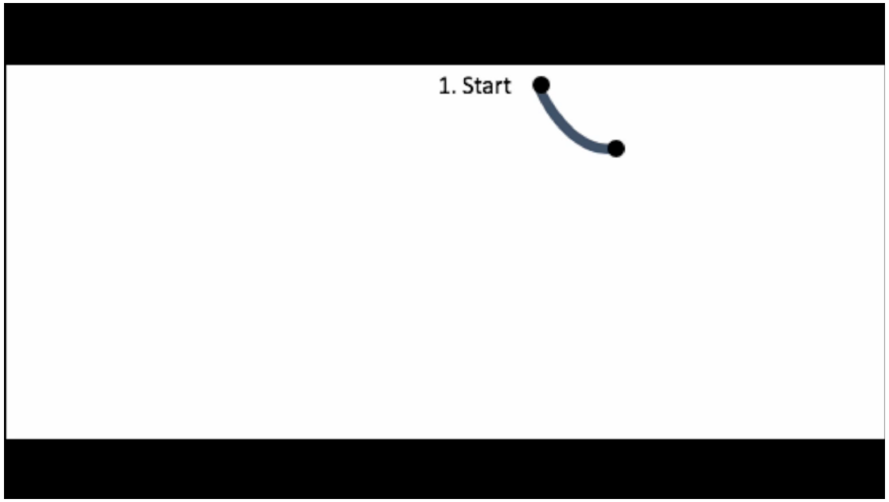


Granularity	Hierarchy	Example
5	Surgical Procedure	Partial Nephrectomy
4	Steps in the procedure	Reconstruction of kidney
3	Tasks in the procedure	Suture
2	Sub-tasks	Pull suture through
1	Motion	Reach and orient

Pick and Place Trajectory and Segments

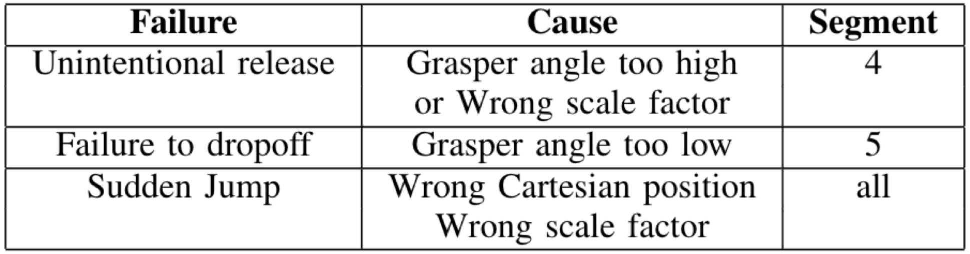


Tracking Surgical Trajectory

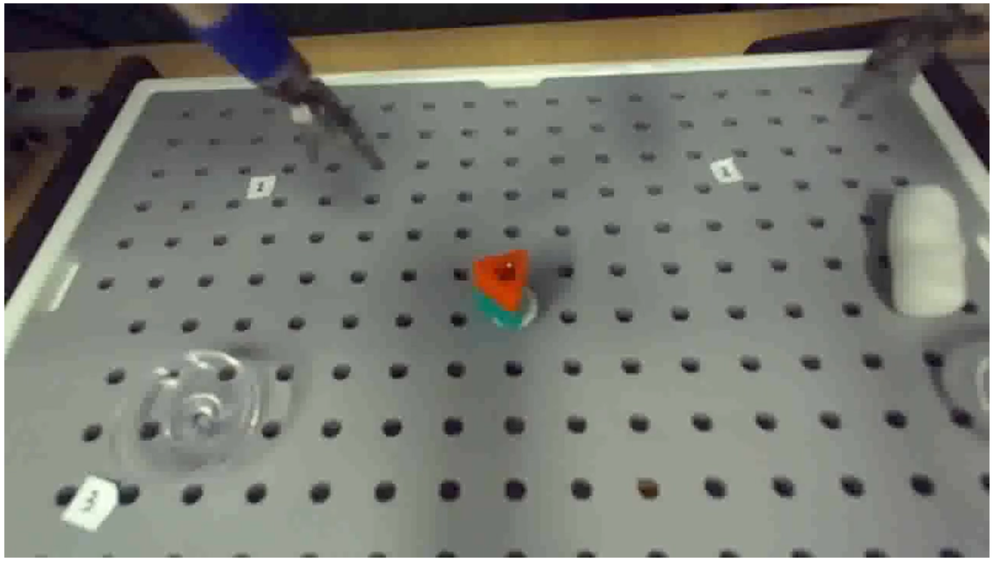


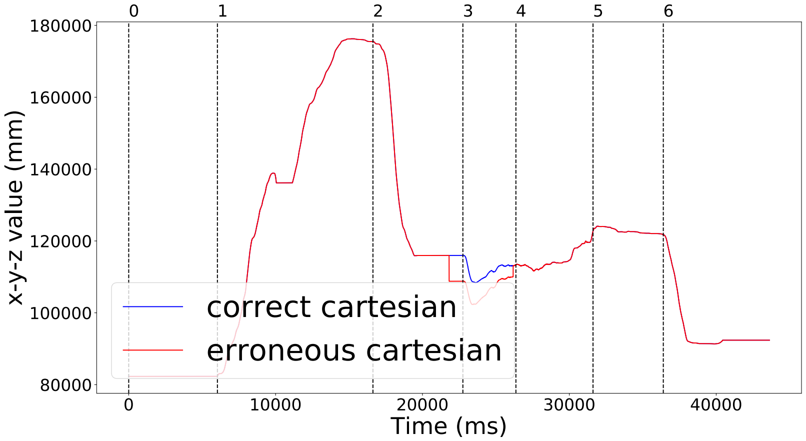
Mapping Trajectory into Segments

Failure Modes in Pick and Place Task



Failure	Cause	Segment
Unintentional release	Grasper angle too high or Wrong scale factor	4
Failure to dropoff	Grasper angle too low	5
Sudden Jump	Wrong Cartesian position Wrong scale factor	all

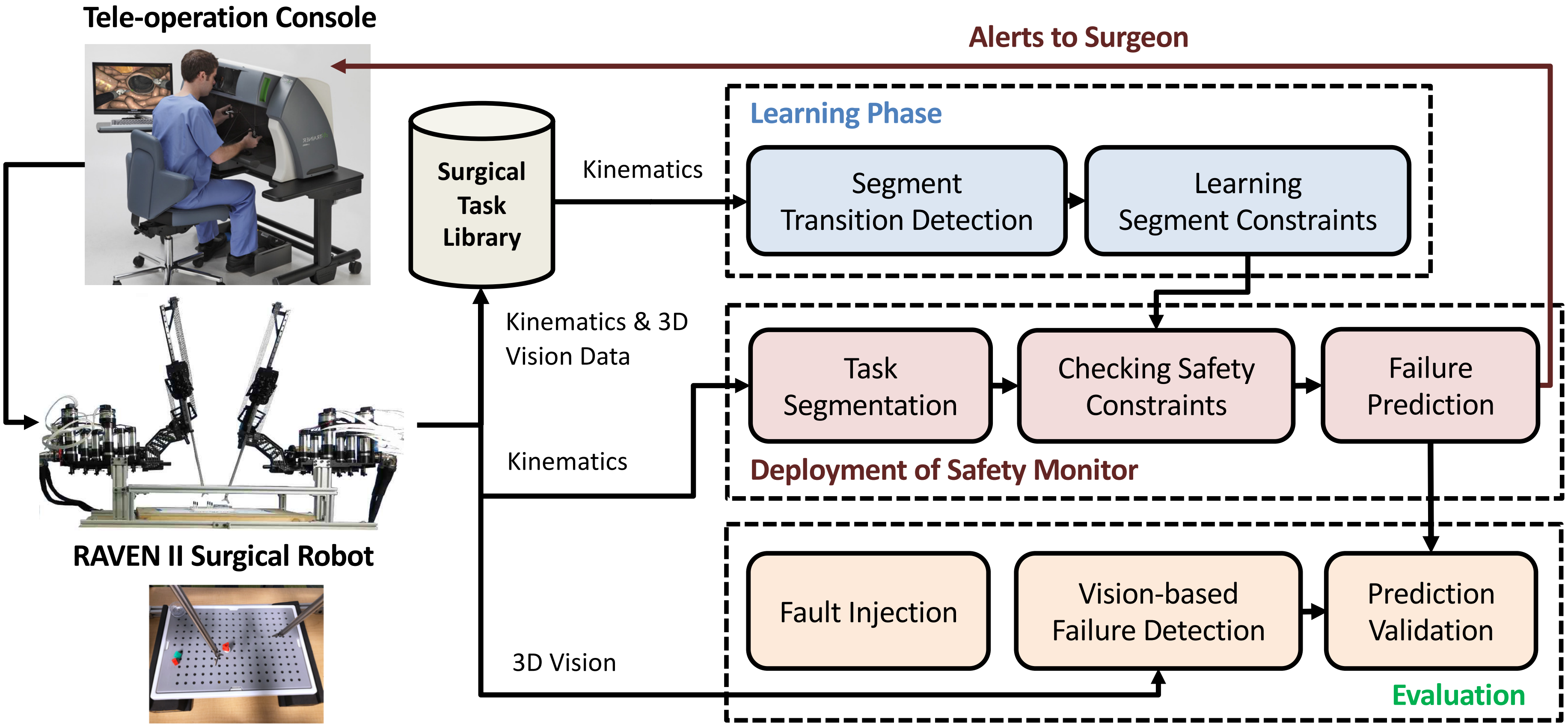




Our Solution

- Actively monitor the movement of the end-effector during fault-free demonstrations of a task
- Learn safety constraints to represent these fault-free behaviors
- Alert the surgeon if we detect a violation of the safety constraints

Context Aware Monitoring, Feedback, Control



Tele-operation Console

Alerts to Surgeon

Surgical Task Library

Kinematics

Learning Phase

Segment Transition Detection

Learning Segment Constraints

Kinematics & 3D Vision Data

Task Segmentation

Checking Safety Constraints

Failure Prediction

Kinematics

Deployment of Safety Monitor

RAVEN II Surgical Robot

Fault Injection

Vision-based Failure Detection

Prediction Validation

3D Vision

Evaluation

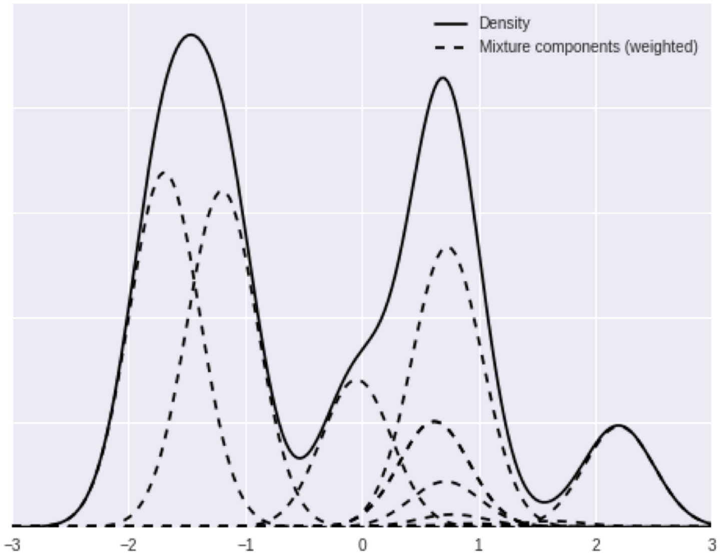
Yasar et al., ISMR 2019.

Detecting Transitions in Segments

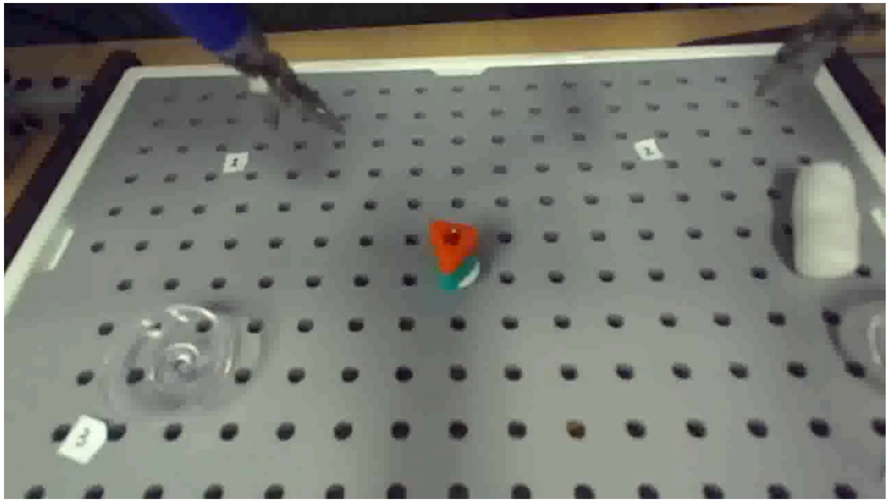
- Find the transitions between the subtasks using unsupervised **Gaussian Mixture Models (GMM)**

- Prior: Number of clusters
- Input to the GMM: Robotic Joint Kinematics values (e.g., position and velocity)

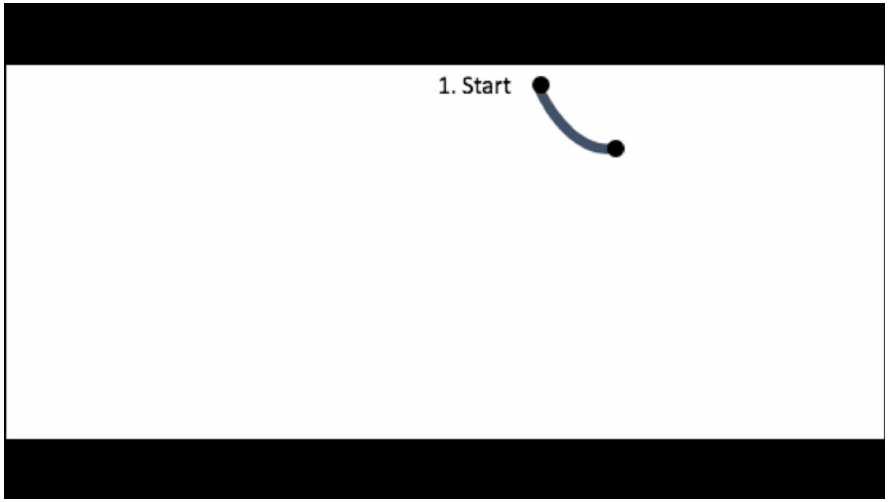
- Previous work (Krishnan et al. 2017) used both kinematics and vision information for detecting segments



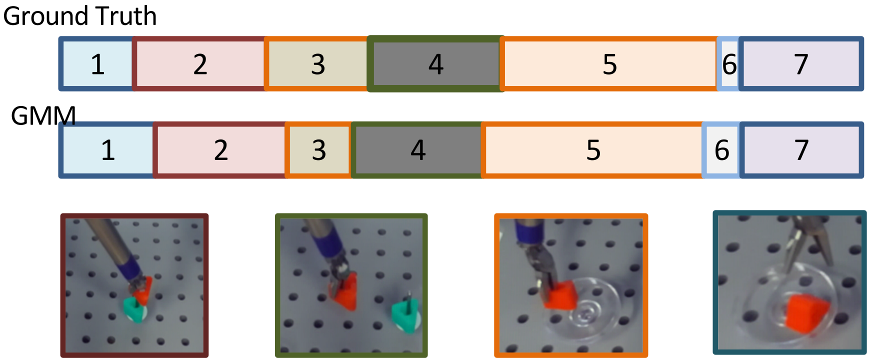
Automated Inference of Context



Tracking Surgical Trajectory



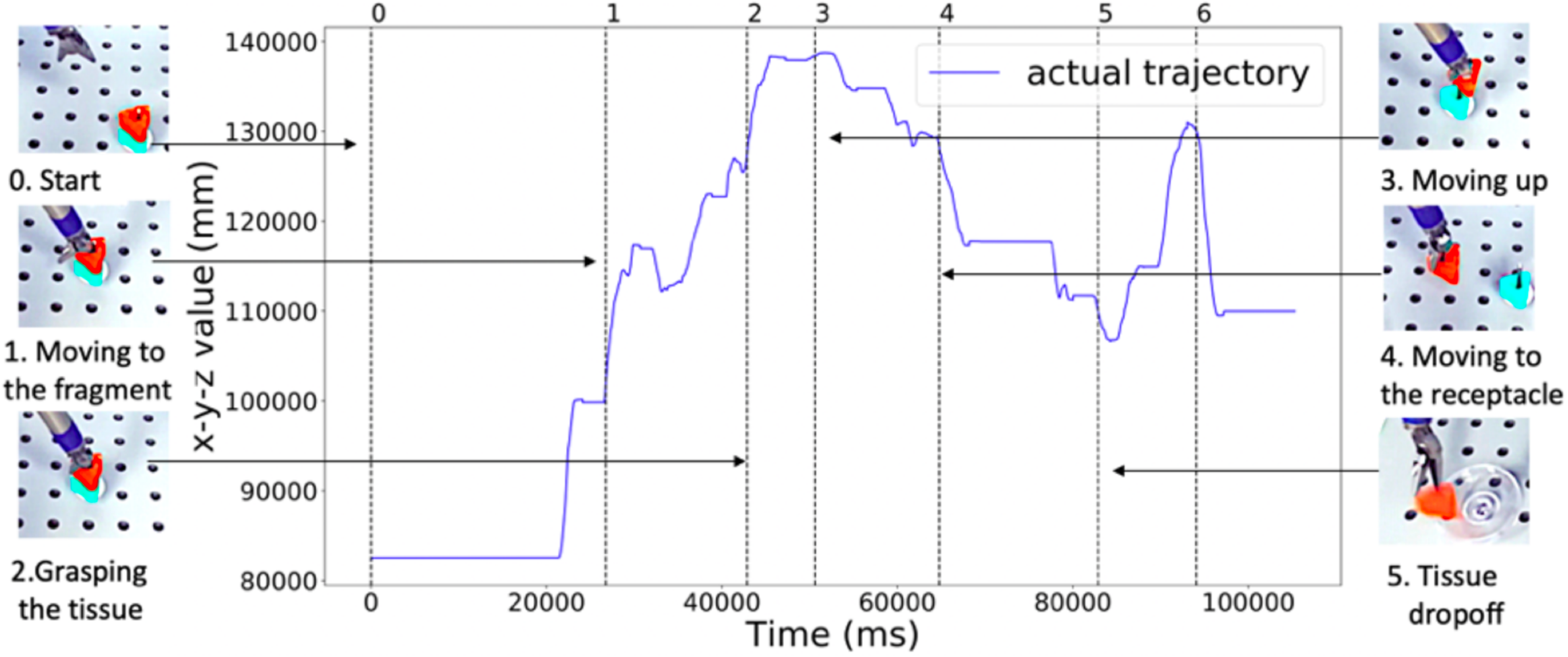
Mapping Trajectory into Segments



Segment 3 Segment 4 Segment 5 Segment 6

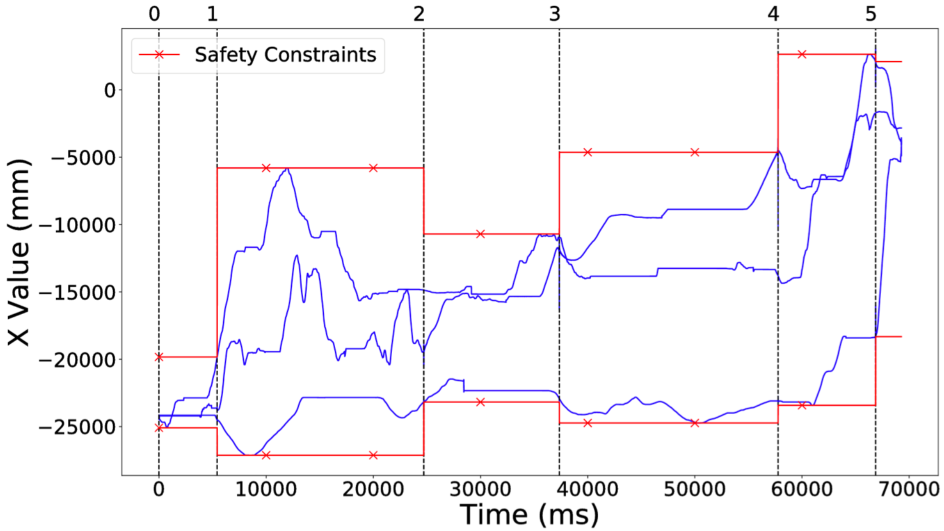
Subtask	Name	Avg. Error (Δt in frames)
0	Start	-56
1	Moving to the block	76
2	Grabbing the block	-69
3	Moving up	-30
4	Moving to the receptacle	-10
5	Dropping the block	-3
6	End	0

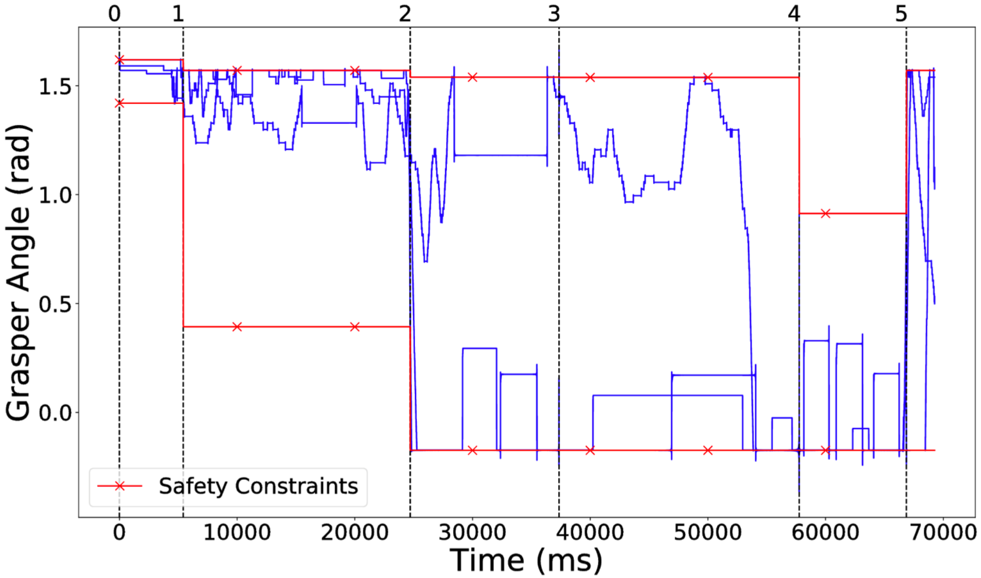
Segment-Specific Safety Constraints



Yasar et al., ISMR 2019.

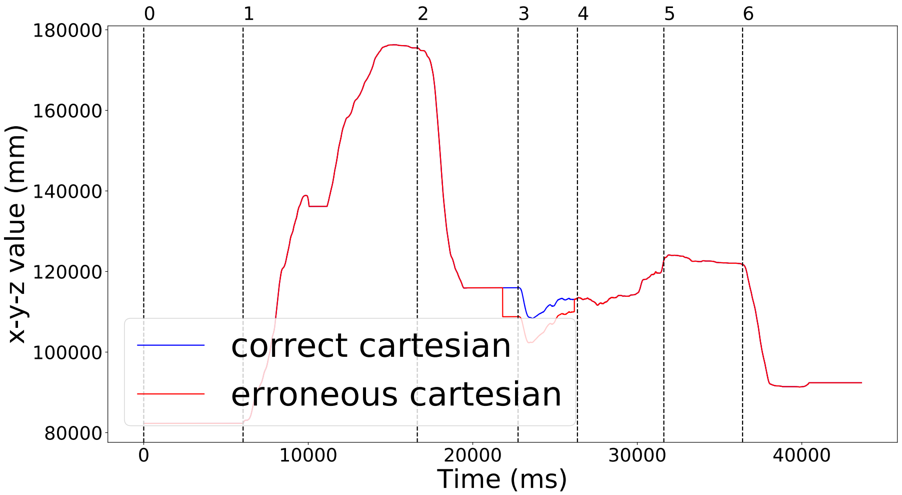
Segment-Specific Safety Constraints



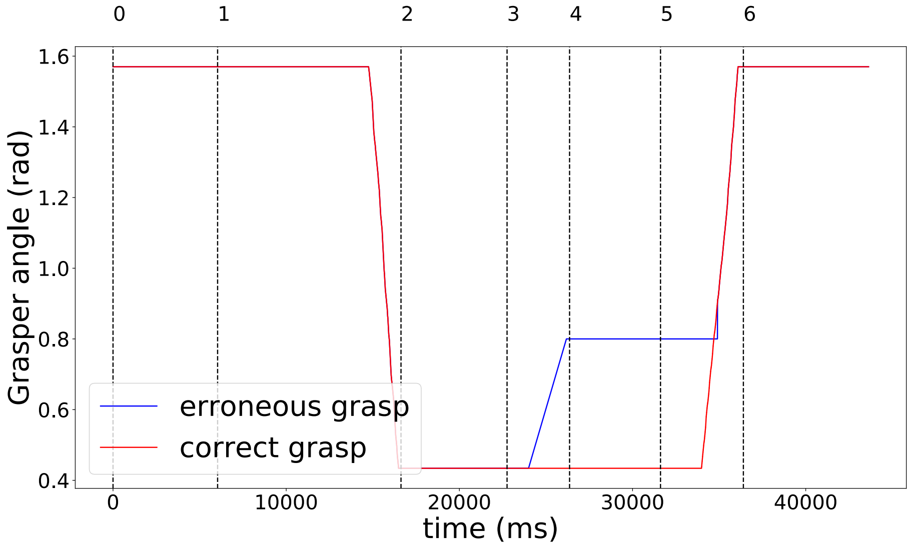


Yasar et al., ISMR 2019.

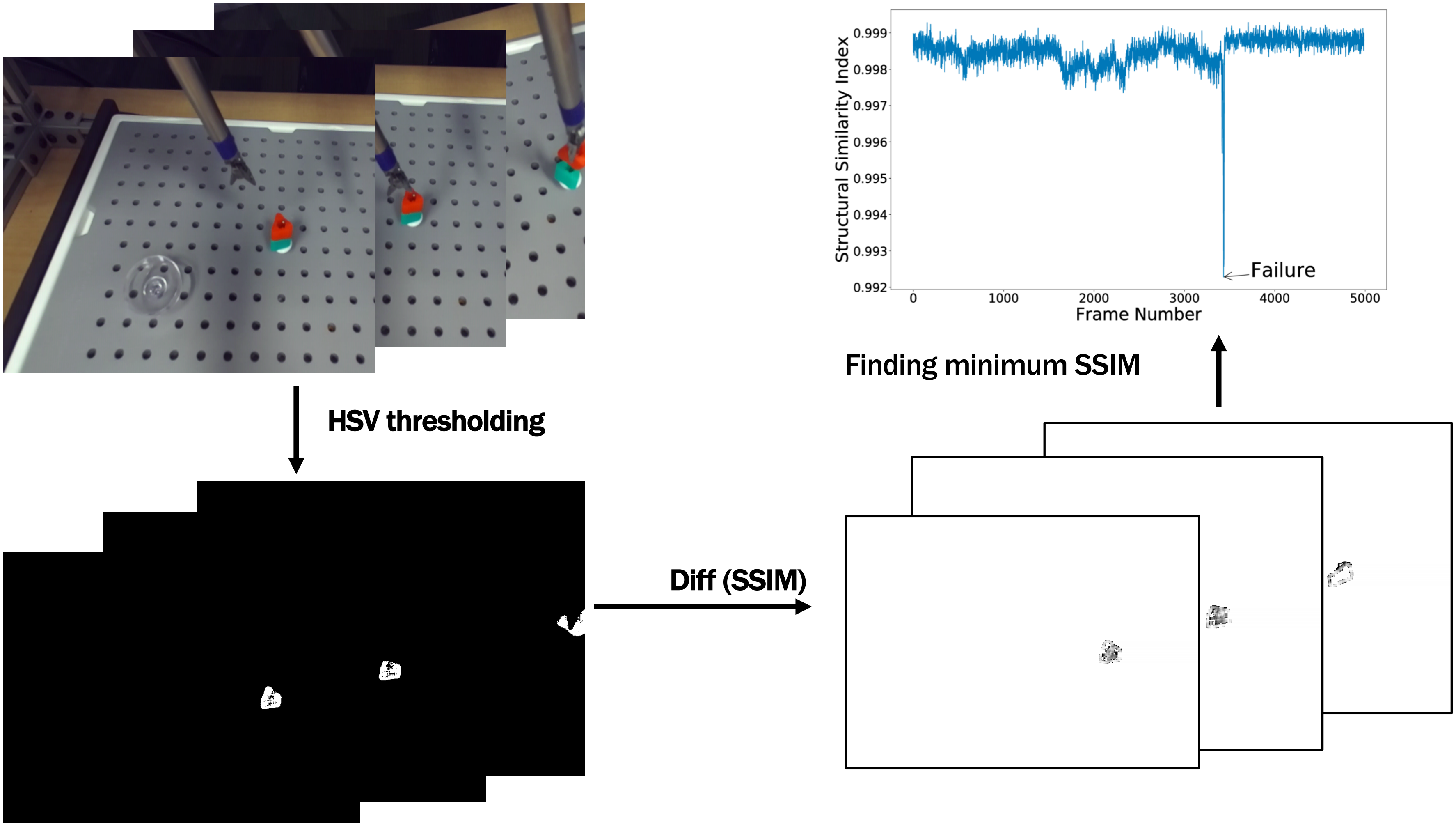
Failure Modes and Fault Injections



Failure	Cause	Segment
Unintentional release	Grasper angle too high or Wrong scale factor	4
Failure to dropoff	Grasper angle too low	5
Sudden Jump	Wrong Cartesian position Wrong scale factor	all



Failure Detection using SSIM

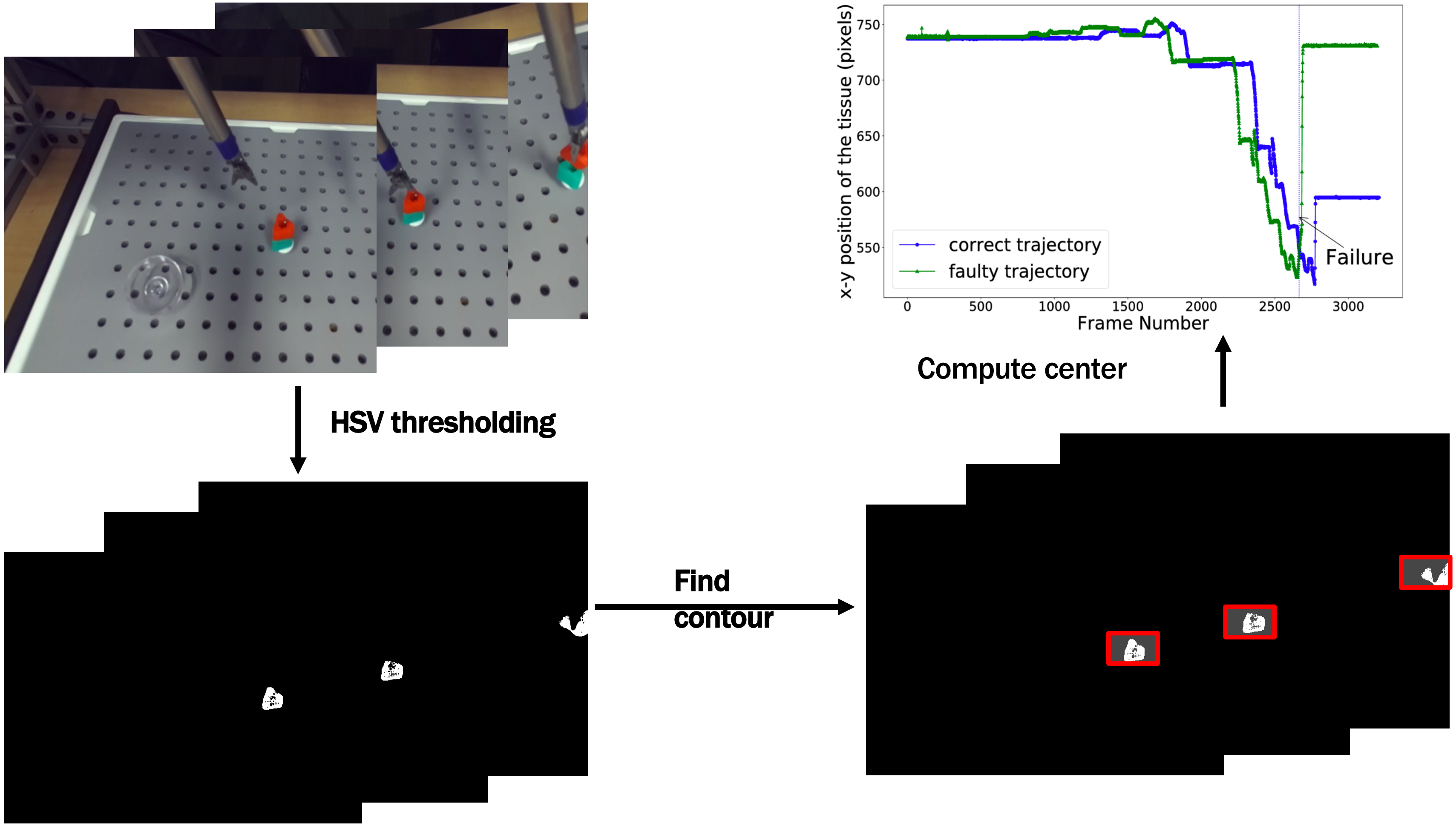


HSV thresholding

Diff (SSIM)

Finding minimum SSIM

Failure Detection via Dynamic Time Warping

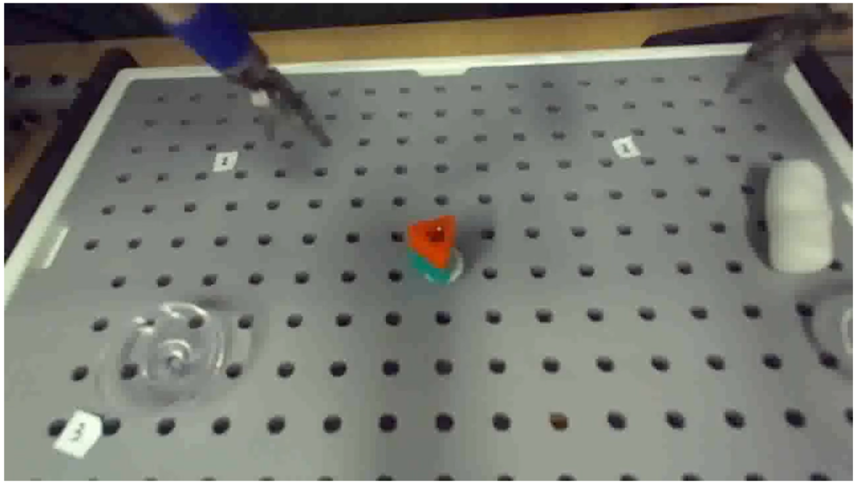


HSV thresholding

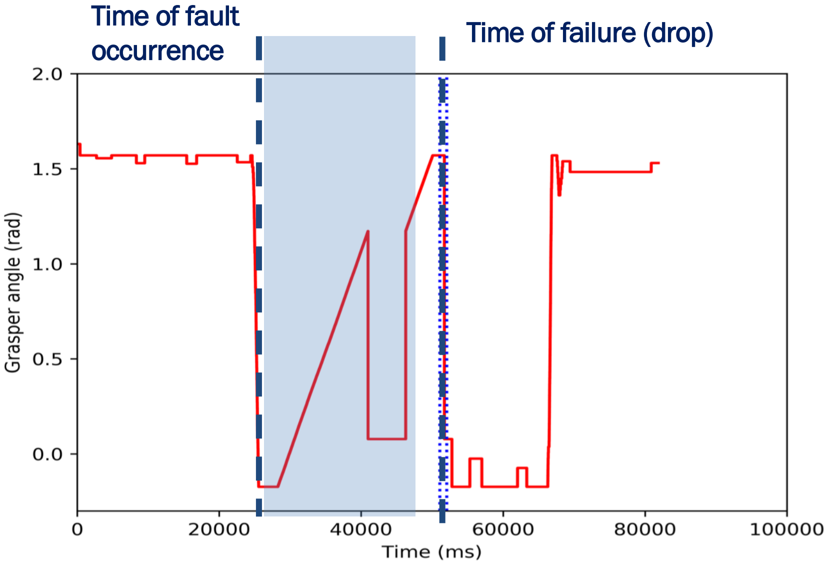
Find contour

Compute center

Early Detection of Safety-Critical Events



Checking Segment Specific Constraints



Time of fault occurrence

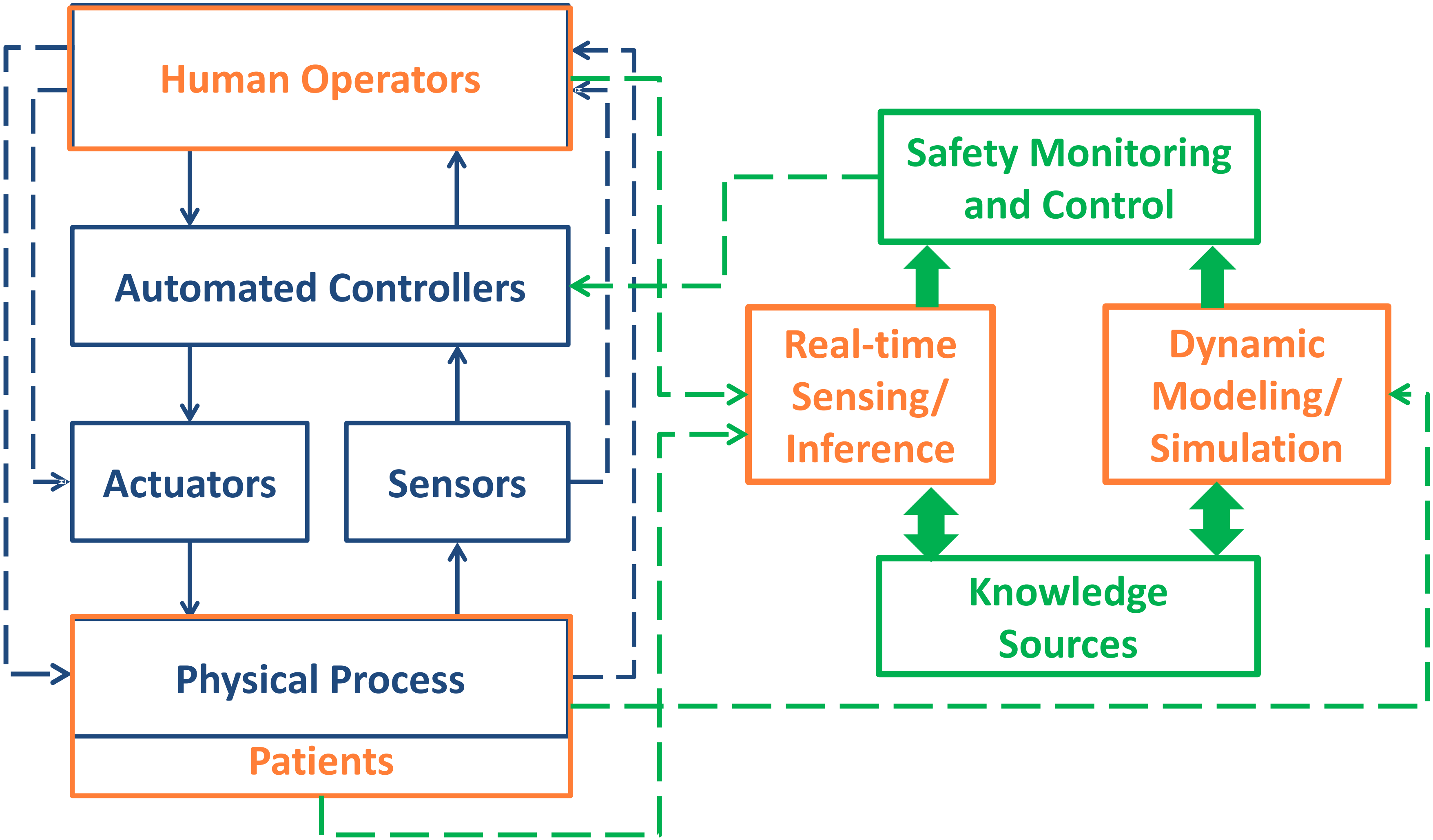
Time of failure (drop)

Detection Window

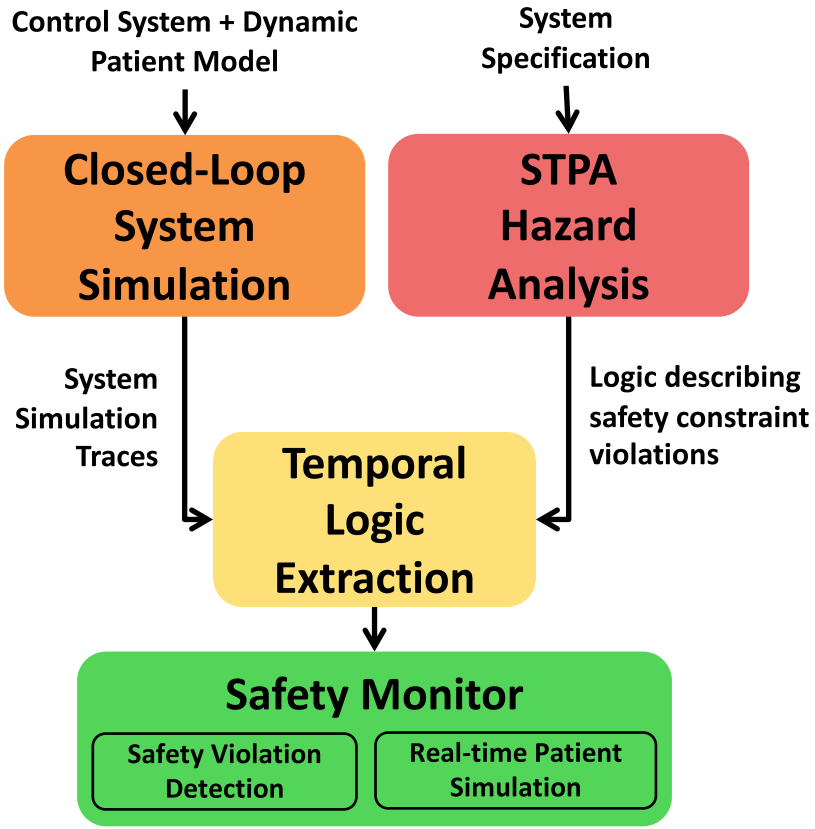
	Overall	Simulation	Dry Lab
Number of Experiments	518	468	50
True Positives	419	398	21
False Positives	95	70	25
True Negatives	3	0	3
False Negatives	1	0	1
Accuracy	**80.8%**		
False Negative Rate	**0.24%**		
False Positive Rate	**96.9%**		

Simulated Failure	**Average Reaction Time**
Sudden Jump	1.7s
Block Drop	14.4s

Context-Aware Safety Monitoring



Automated Synthesis of Context-Aware Safety Monitors



Control System + Dynamic Patient Model

System Specification

Closed-Loop System Simulation

STPA Hazard Analysis

System Simulation Traces

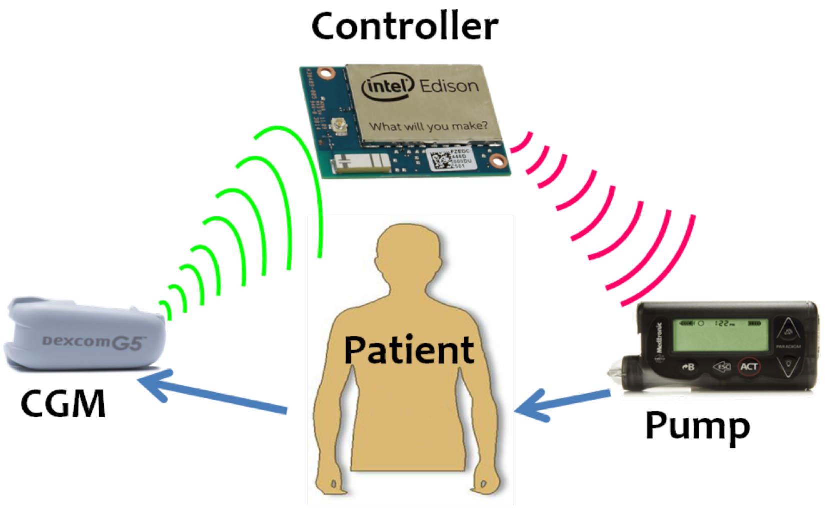
Logic describing safety constraint violations

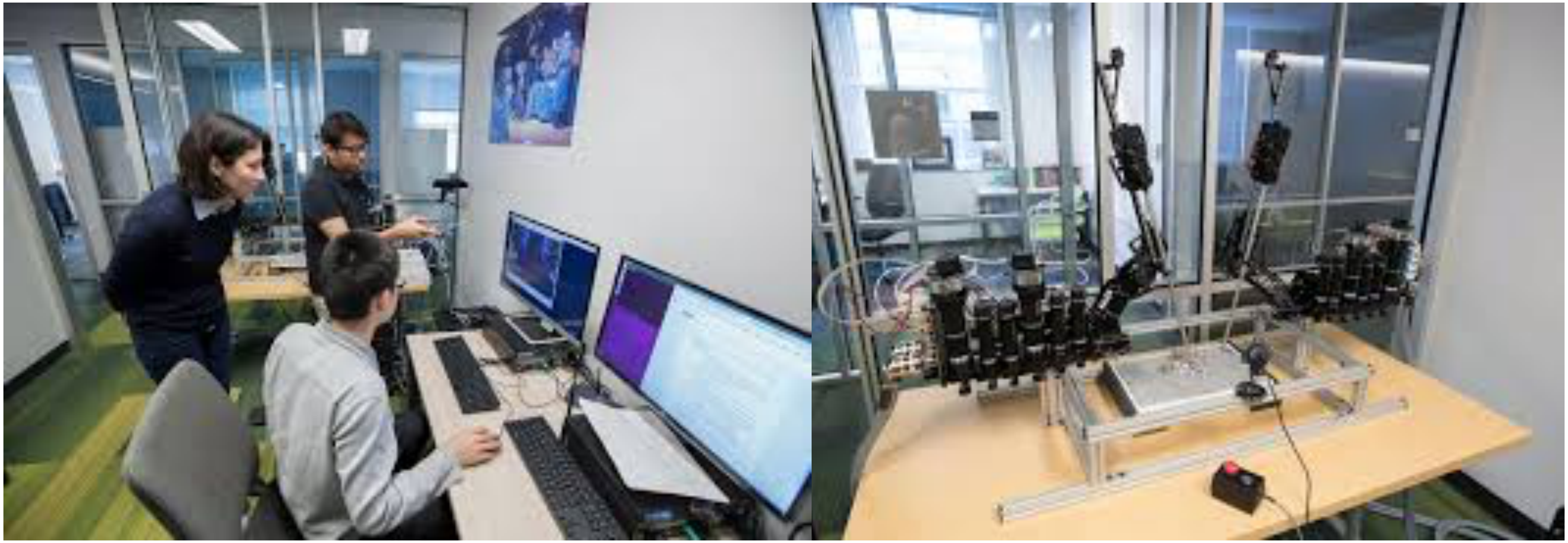
Temporal Logic Extraction

Safety Monitor

Safety Violation Detection

Real-time Patient Simulation





UVA Dependable Systems and Analytics Group

NSF

NIST National Institute of Standards and Technology U.S. Department of Commerce

UNIVERSITY of VIRGINIA ENGINEERING

UNIVERSITY of VIRGINIA HEALTH SYSTEM