NRI: INT: COLLAB: Manufacturing USA: Intelligent Human-Robot Collaboration for Smart Factory

- Robert X. Gao, Department of Mechanical & Aerospace Engineering, Case Western Reserve University
- Ming Leu, Department of Mechanical & Aerospace Engineering, Missouri University of Science and Technology
- Gloria Wiens, Department of Mechanical & Aerospace Engineering, University of Florida
- Zhaozheng Yin, AI Institute, Department of Biomedical Informatics and Department of Computer Science, Stony Brook University

Partners:

COMAU

STIIMA

National Research Council of Italy

Manufacturing USA

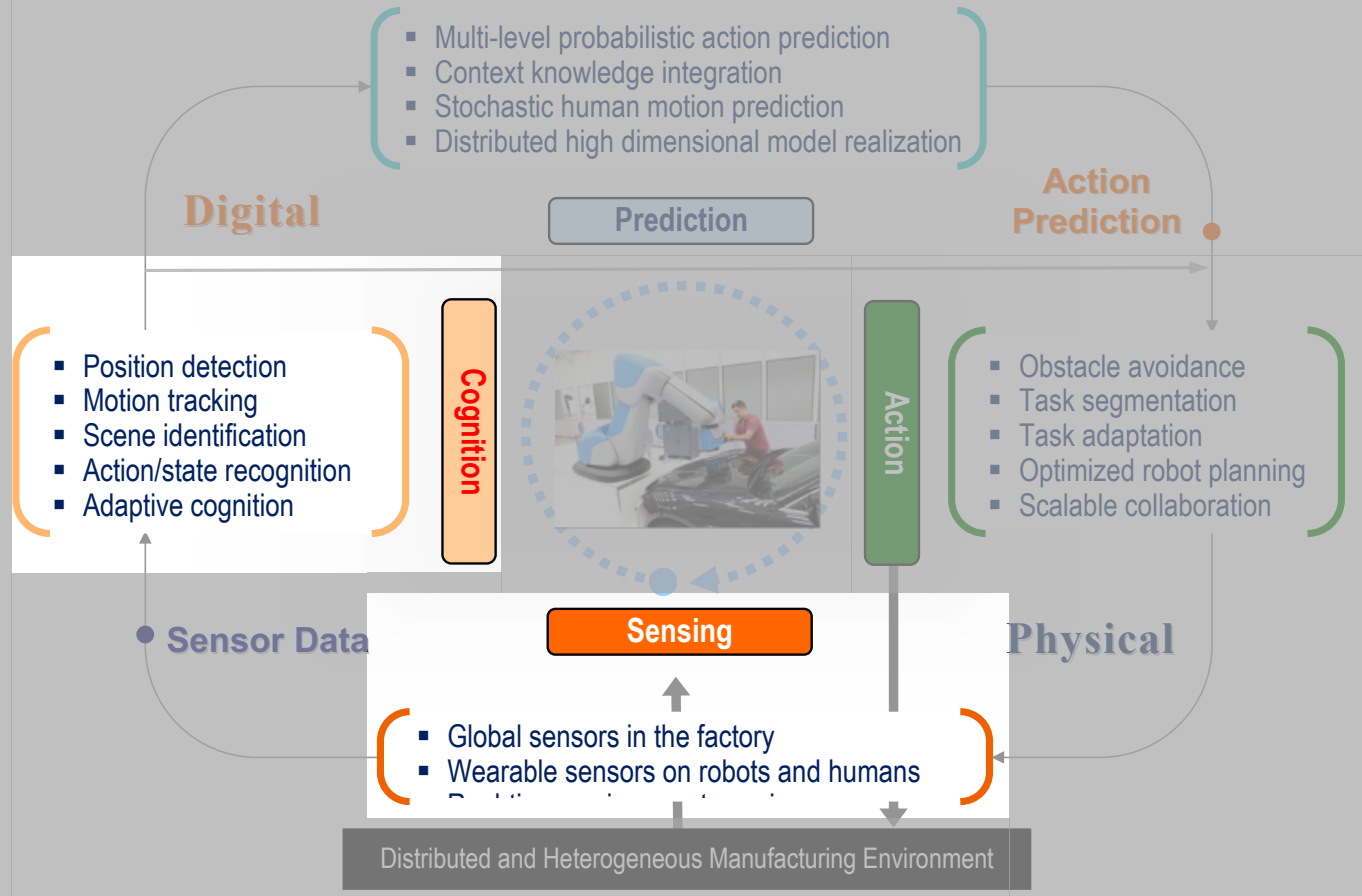
NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

NRI: INT: COLLAB: Manufacturing USA: Intelligent Human-Robot Collaboration for Smart Factory

Research Goals:

- We propose four research components to realize the envisioned human-robot collaboration (HRC) for an automated HRC manufacturing cell,
  - from data acquisition in the physical domain
  - to data manipulation in the digital domain, and
  - back to robot control in the physical domain:

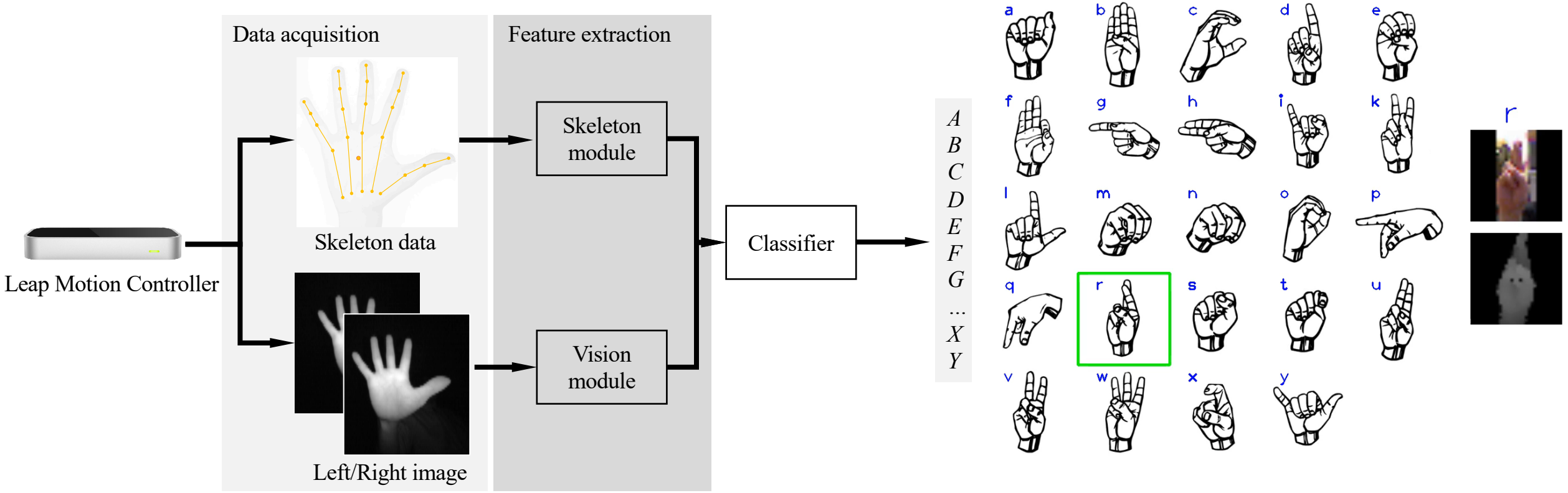
sensing, cognition, prediction & action.



NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

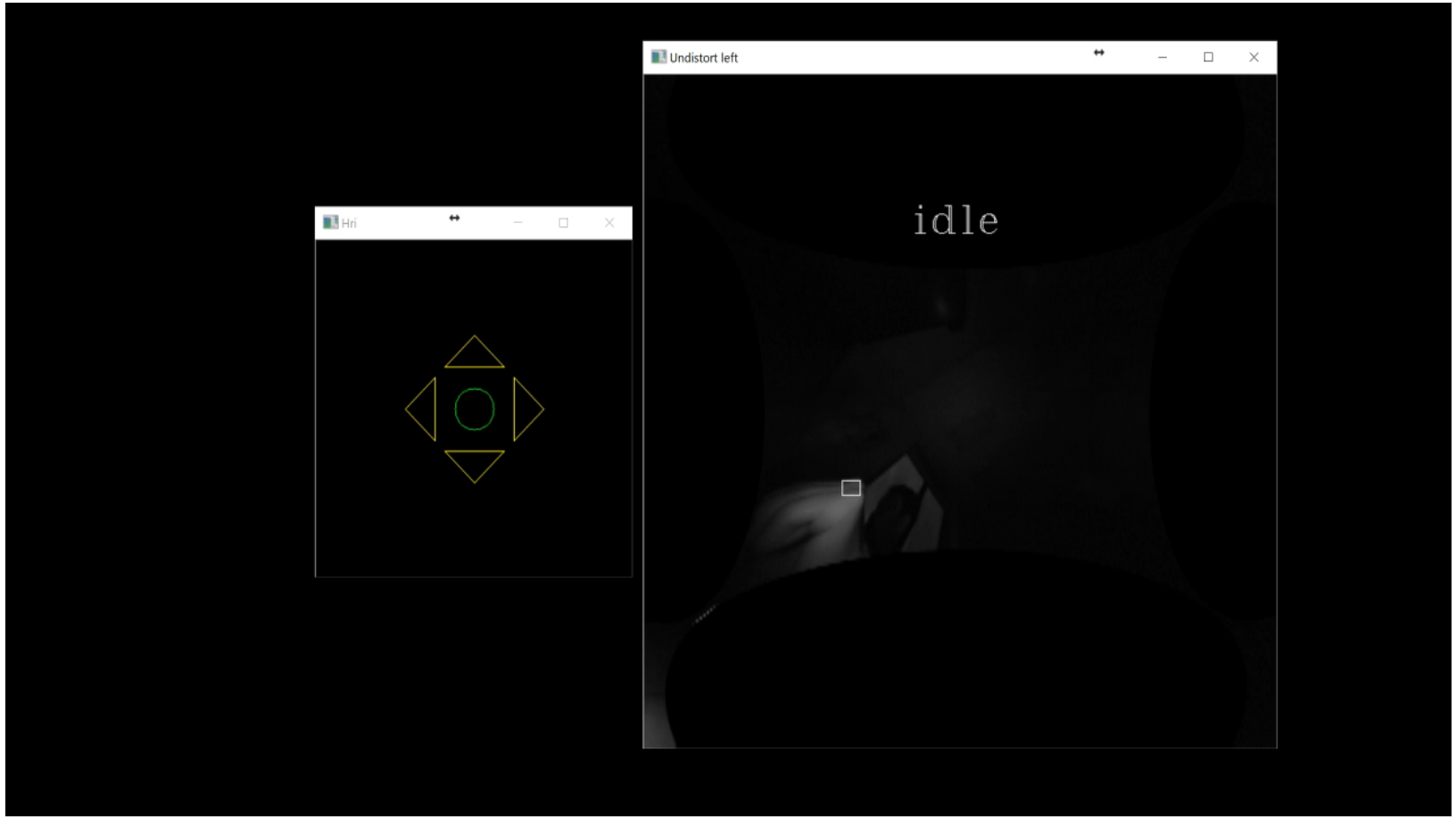
- **Stationary Gesture Recognition**



NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

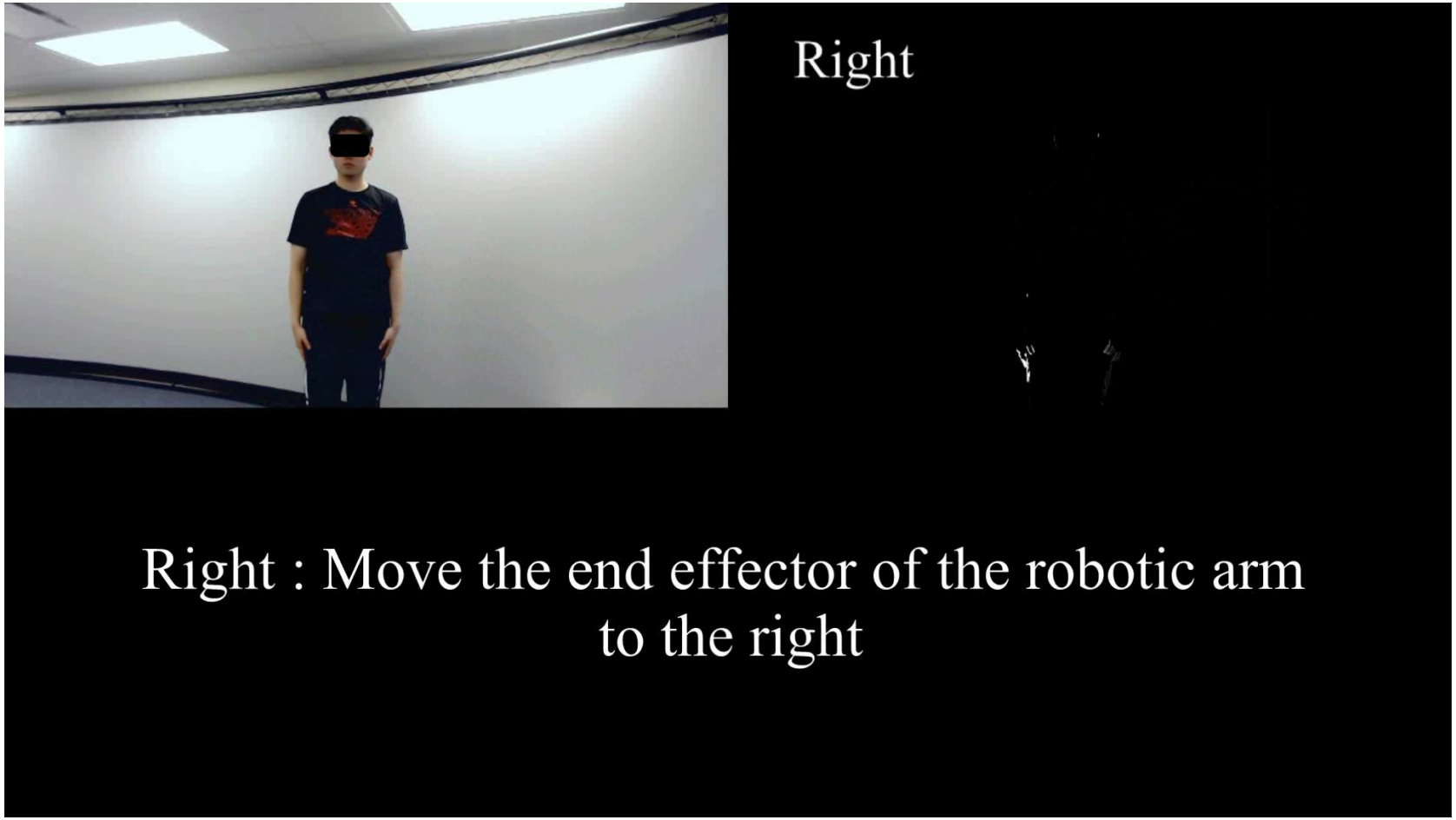
- **Stationary Gesture Recognition**



NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

- **Dynamic Gesture Recognition by Motion History Images**

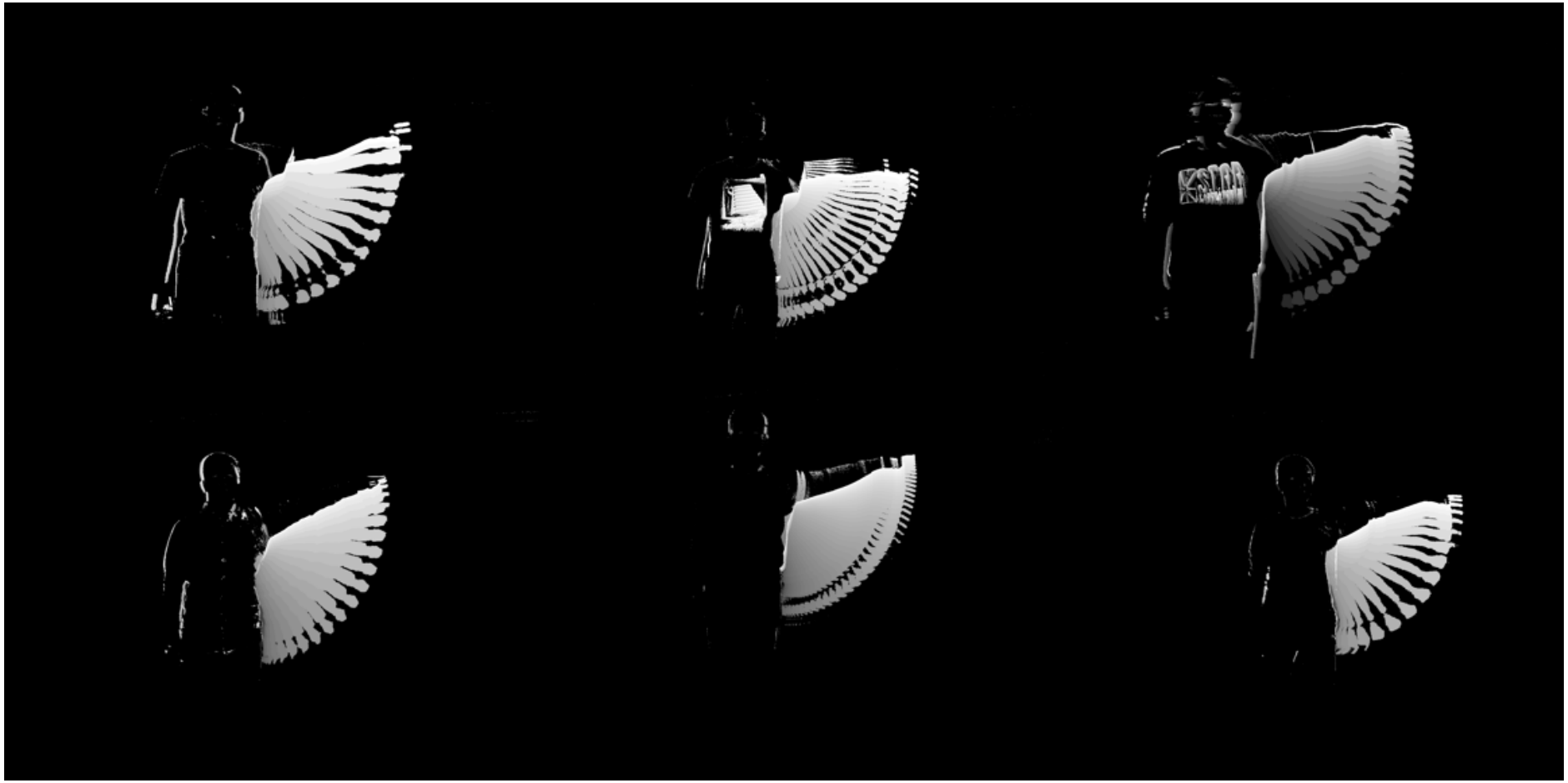


NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

- **Dynamic Gesture Recognition by Motion History Images**

The gesture ‘left’ from six subjects.



NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

- **Dynamic Gesture Recognition by Motion History Images**

(a) start

(b) stop

(c) up

(d) down

(e) left

(f) right

(g) inward

(h) outward

(i) open gripper

(j) close gripper

NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

- **Dynamic Gesture Recognition by Motion History Images**


(a) start


(b) stop


(c) up


(d) down


(e) left


(f) right


(g) inward


(h) outward


(i) open gripper


(j) close gripper

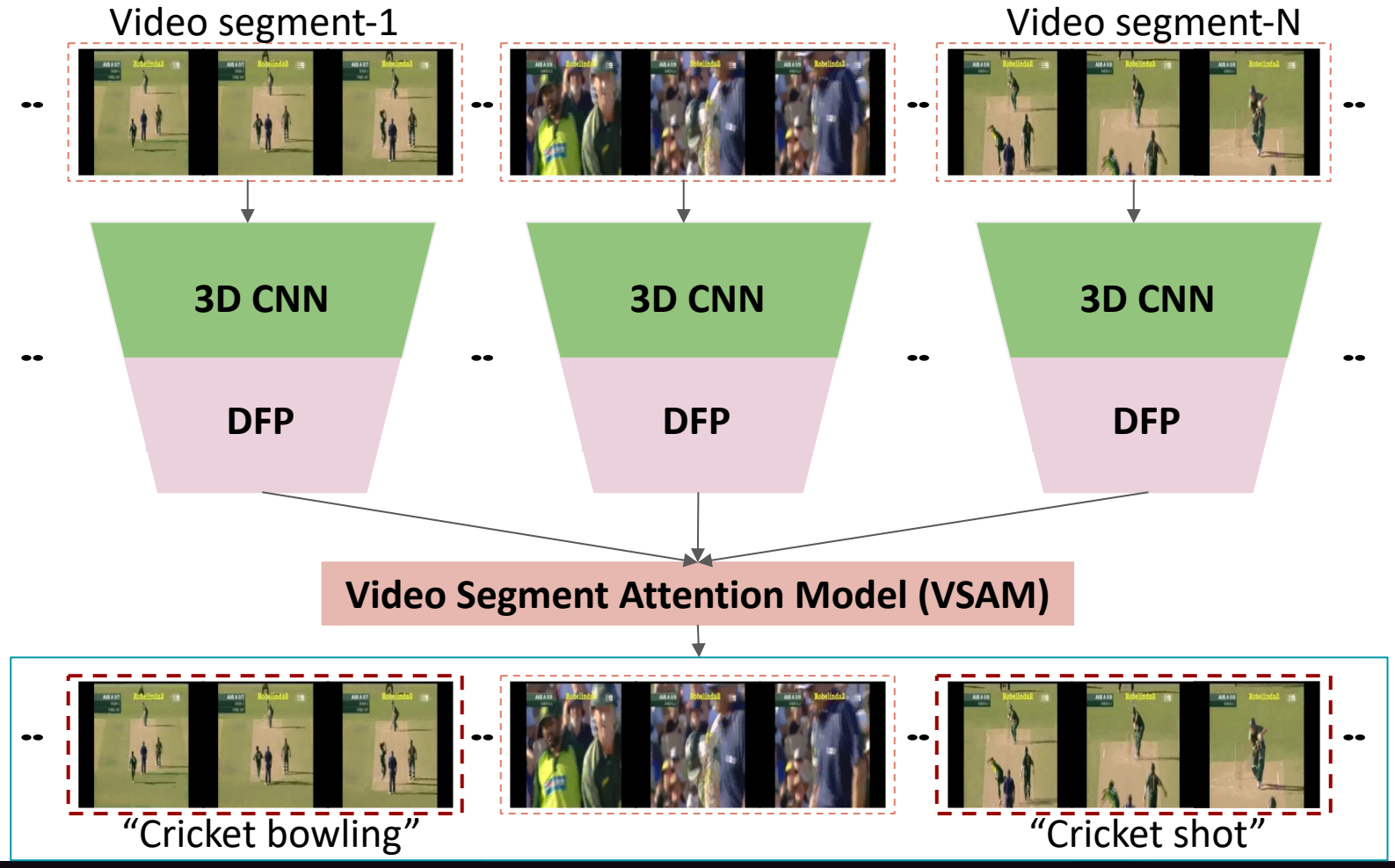
The classification performance of a CNN

Classes	TP	TN	FP	FN	Accuracy	Precision	Recall	F-score
start	93	822	1	0	0.999	0.989	1.000	0.995
stop	102	813	1	0	0.999	0.990	1.000	0.995
up	93	822	0	1	0.999	1.000	0.989	0.995
down	92	823	0	1	0.999	1.000	0.989	0.995
left	88	827	0	0	1.000	1.000	1.000	1.000
right	90	825	0	0	1.000	1.000	1.000	1.000
inward	91	824	0	0	1.000	1.000	1.000	1.000
outward	90	825	0	2	0.998	1.000	0.978	0.989
open	88	827	0	0	1.000	1.000	1.000	1.000
close	88	827	2	0	0.998	0.978	1.000	0.989

NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

- **Action Recognition from Video Clips**

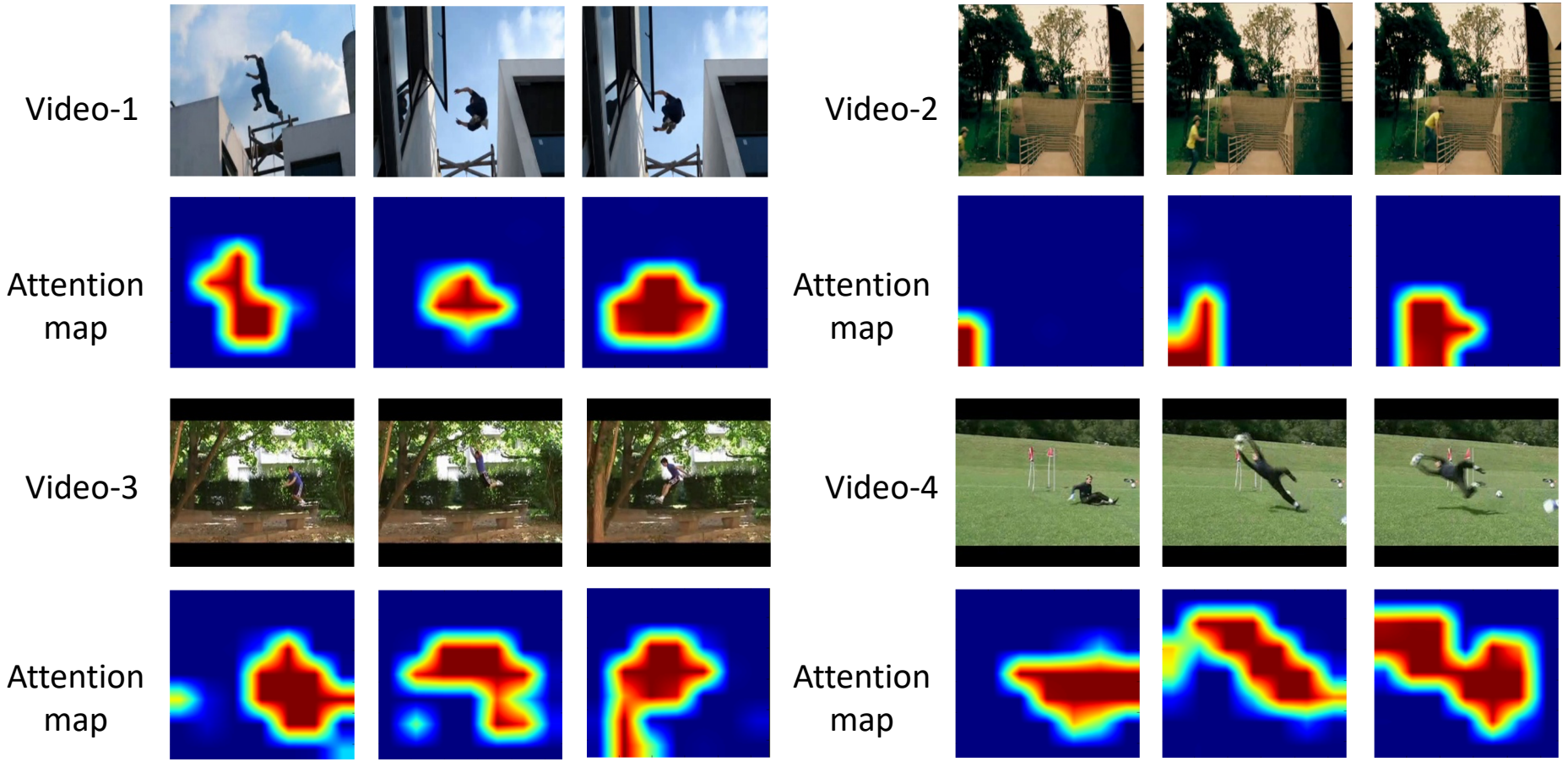


NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

- **Action Recognition from Video Clips**

Qualitative examples on the ‘jump’ action:

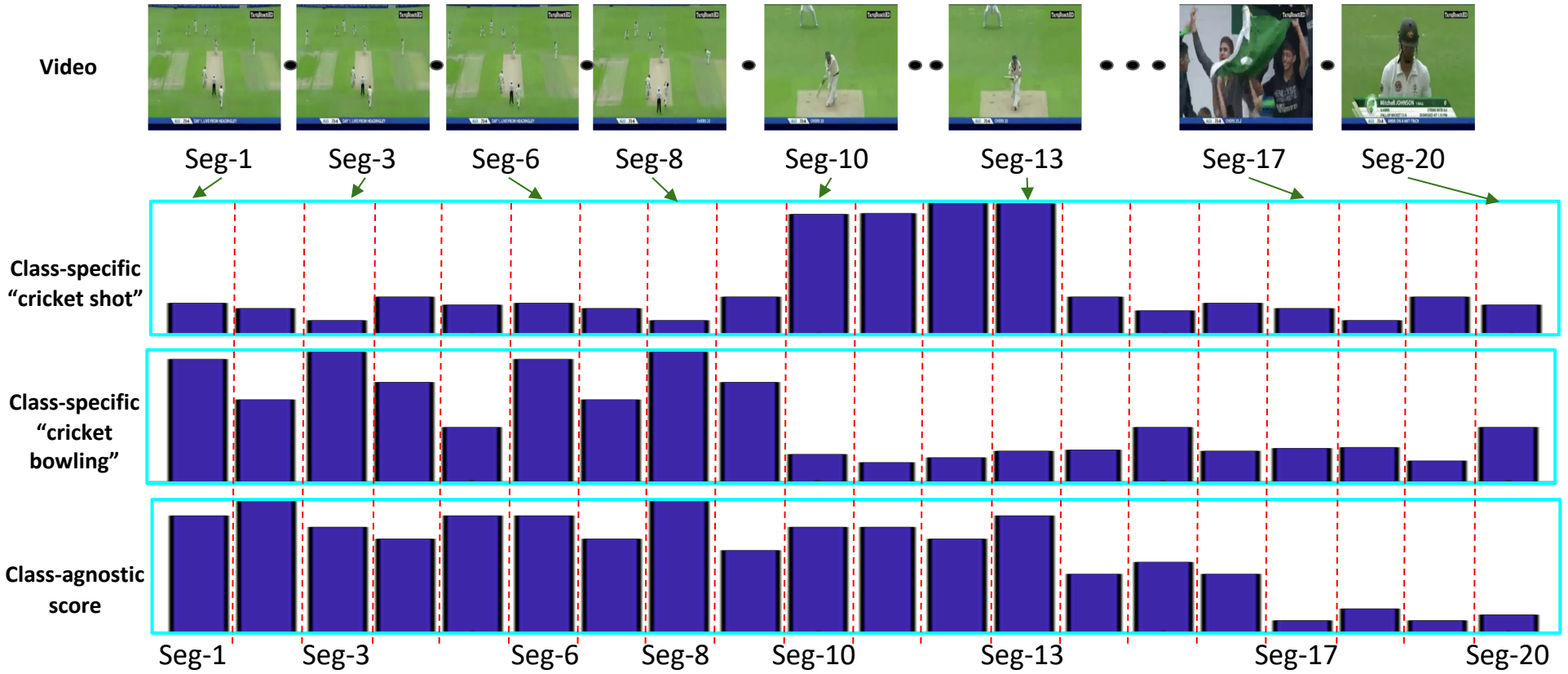


NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

- **Action Recognition from Video Clips**

Qualitative example on the ‘cricket game’ action:



NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

- **Action Recognition from Video Clips**

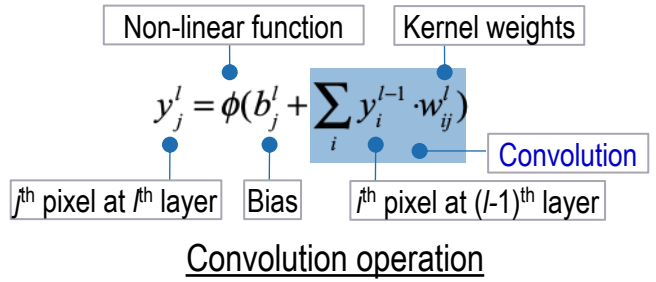
THUMOS (2014) Dataset:

- Containing 101 action classes
- Training dataset (Trimmed) → (13320 videos)
- Validation Dataset (Untrimmed) → (1010 videos)
- Background Dataset (Untrimmed) → (2500 videos)
- Testing Dataset (Untrimmed) → (1574 videos)
  - Containing action frames → 1094 videos
  - Only Background videos → 480 videos

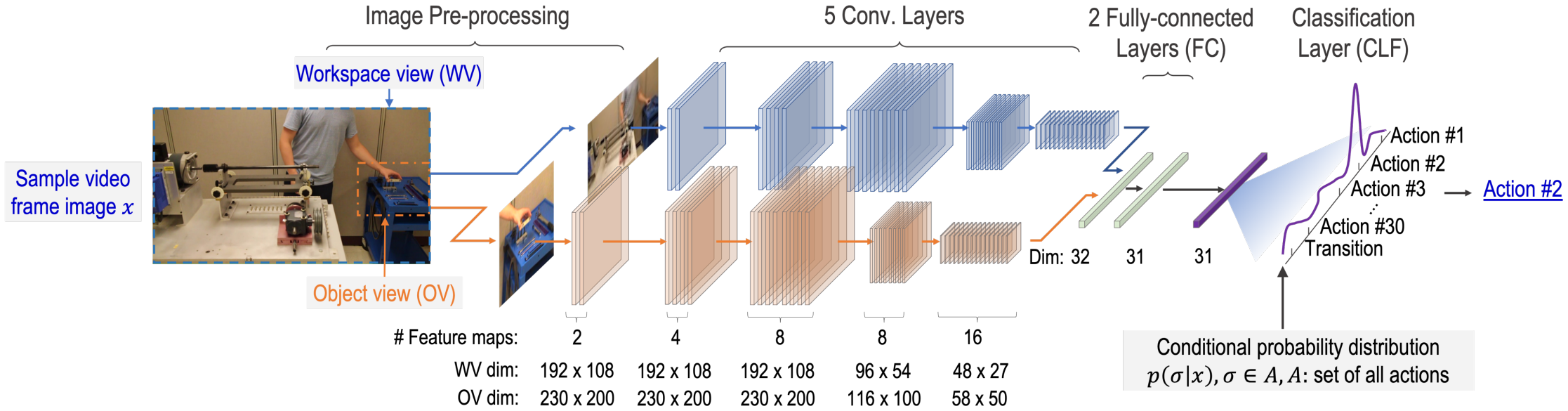
NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing and Cognition

- Bi-stream CNN structure
  - ✓ Each image pre-processed into:
    - ➢ Workspace view: human worker position and posture for action recognition
    - ➢ Object view: for object recognition



Convolution operation



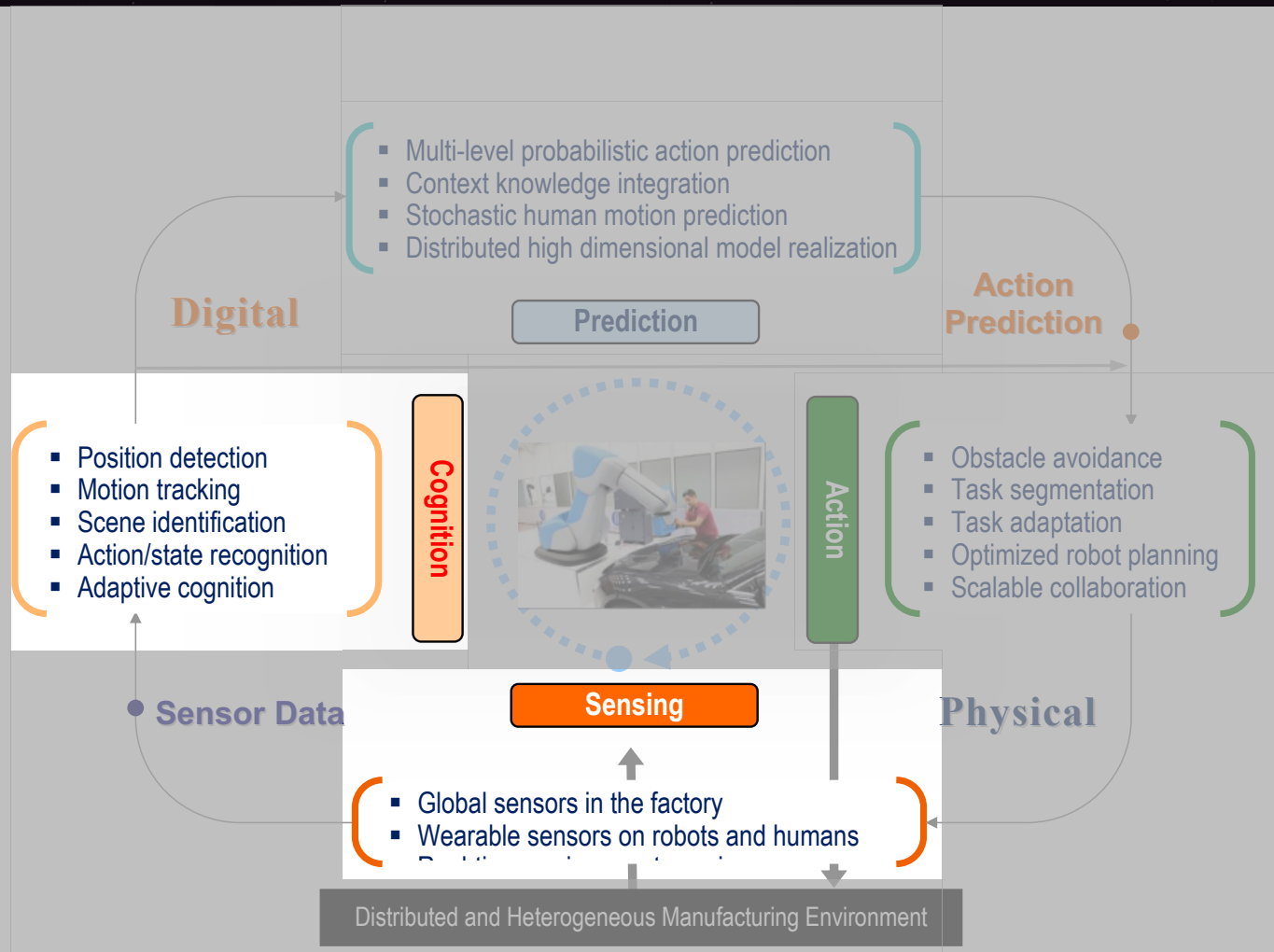
NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

NRI: INT: COLLAB: Manufacturing USA: Intelligent Human-Robot Collaboration for Smart Factory

Research Goals:

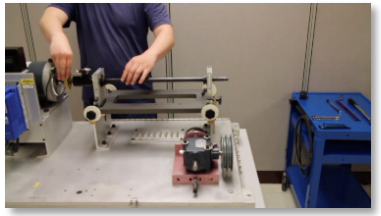
- We propose four research components to realize the envisioned human-robot collaboration (HRC) for an automated HRC manufacturing cell,
  - from data acquisition in the physical domain
  - to data manipulation in the digital domain, and
  - back to robot control in the physical domain:

sensing, cognition, prediction & action.



NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Sensing, Cognition, and Prediction

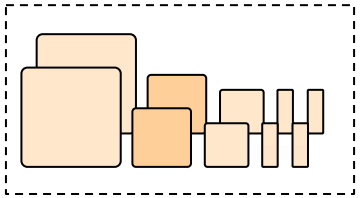


Workspace

Sensing

Video recording

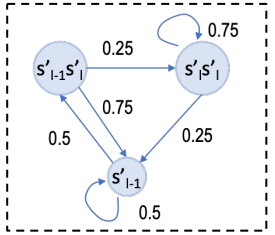
Action Recognition



Deep Convolutional Neural Networks

Sequence of current and past actions (e.g. $s'=s'_1,s'_2,\ldots,s'_l$)

Action Prediction



Variable-order Markov Model

Predicted future action

$\arg\max_w p(w \mid s'_{l-L+1},\ldots,s'_l)$

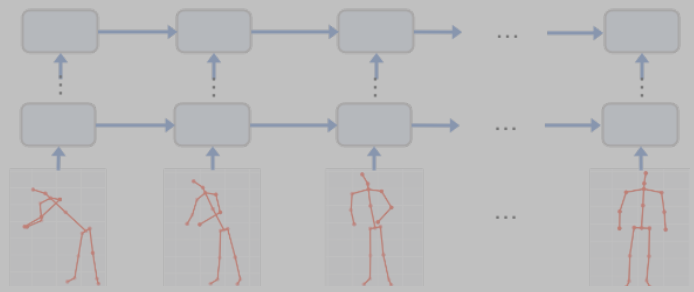
Pass *which* tool to worker?

Action

Pose tracking

Historic trajectory of human pose

Trajectory Prediction



Recurrent Neural Networks

Predicted future human pose

Where to pass the tool?

NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Action Prediction

Variable-order Markov Model

- ✓ Establish *action transition graph* and predict future action via transition probability conditioned on *variable length* of current and past actions
- ✓ *Optimal context* determined as specific current and past action sequence that maximizes transition probability

- a Pick up screw
- b Install screw
- c Pick up cap
- d Install cap
- e Pick up/return Allen key
- f Turn Allen key

Workspace video

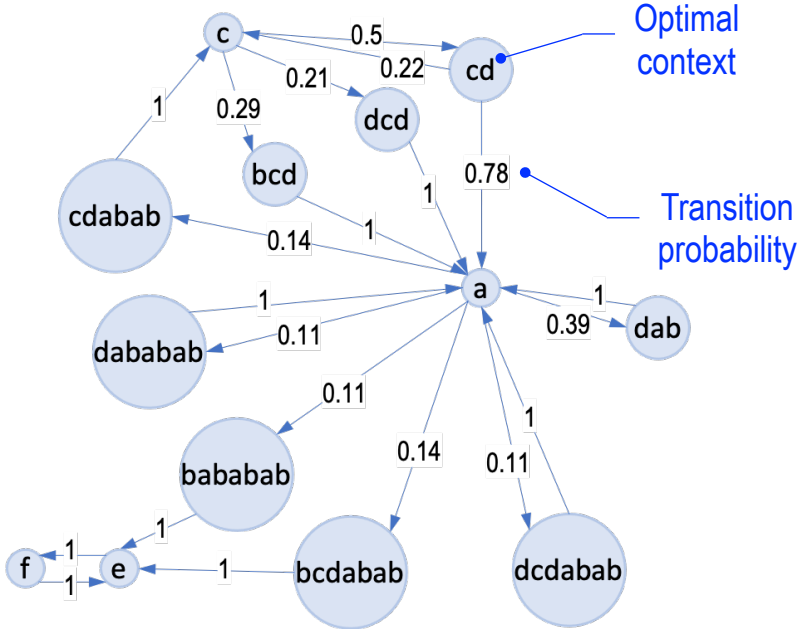
(c, d, c, d, a, b, a, b, a, b, a, b, e, f, e)

(c, d, a, b, a, b, c, d, a, b, a, b, e, f, e)

(c, d, a, b, a, b, e, f, e, c, d, a, b, a, b, e, f, e)

⋮

Recognized sequences
(via DCNN)

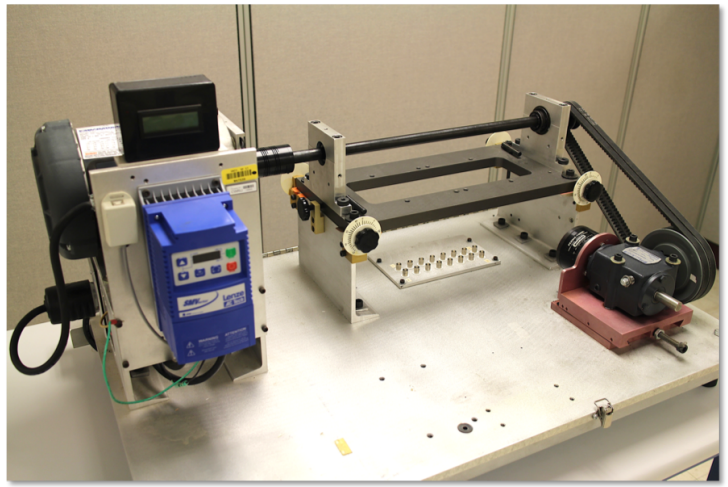


Action transition graph

NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

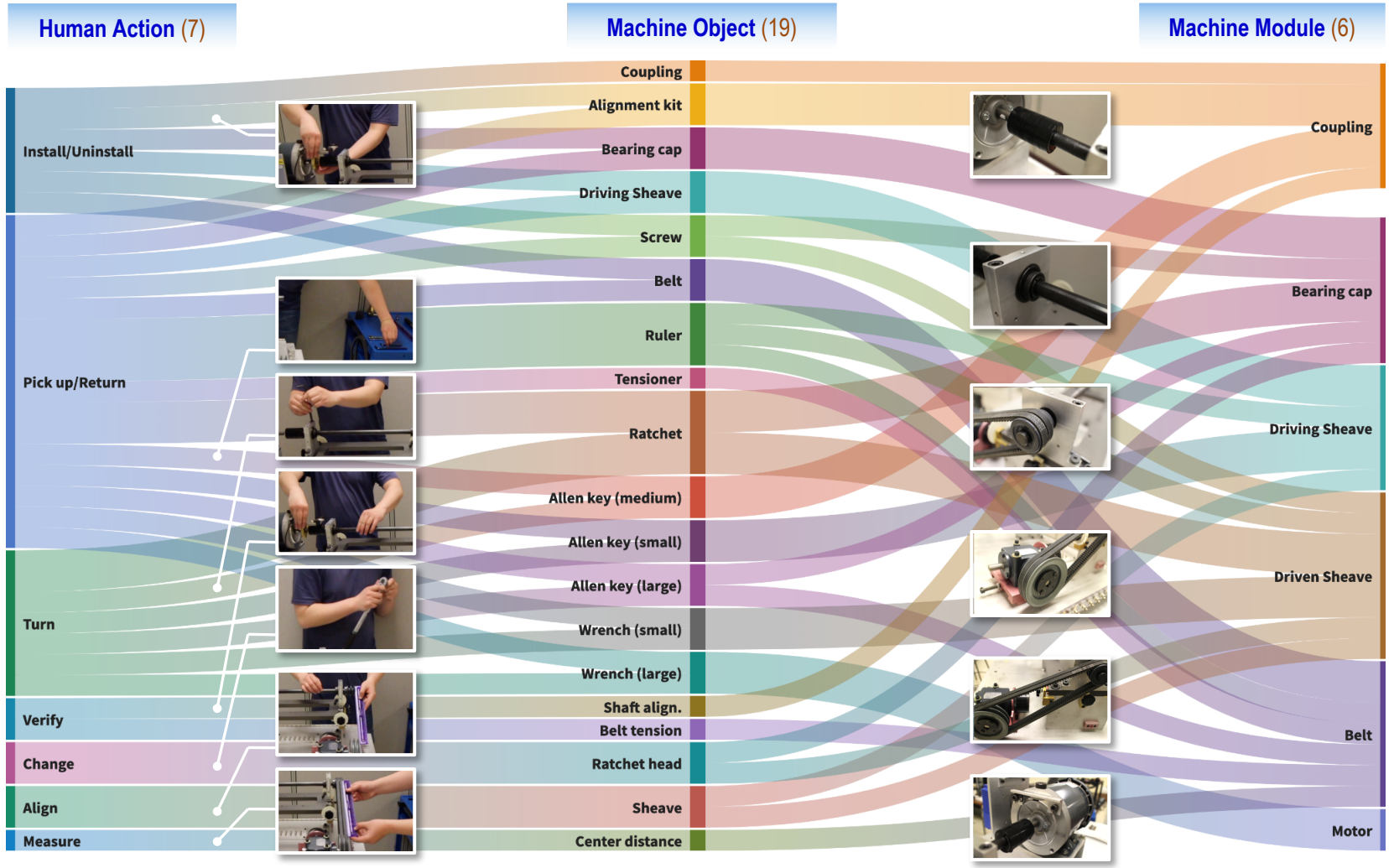
Experimental Evaluation: Action Prediction (1)

Assembly Test Platform



Machinery Fault Simulator (MSF) Testbed
SpectraQuest, Inc.

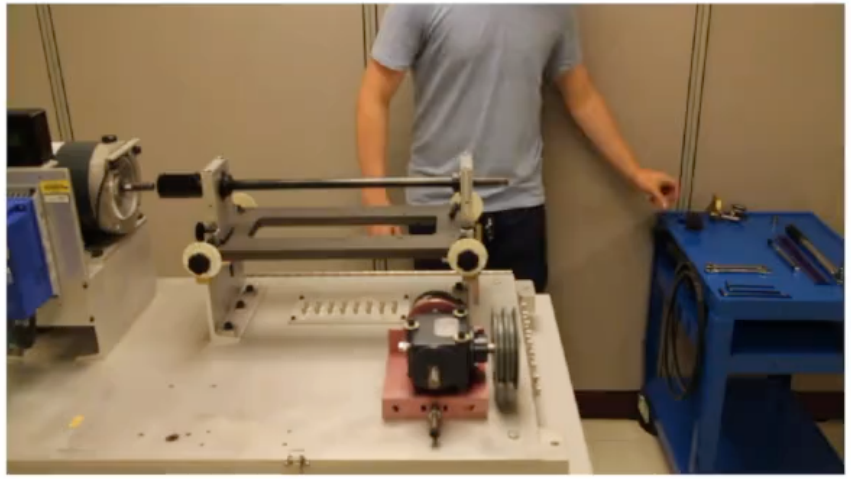
30 unique *atomic* action/object pairs identified



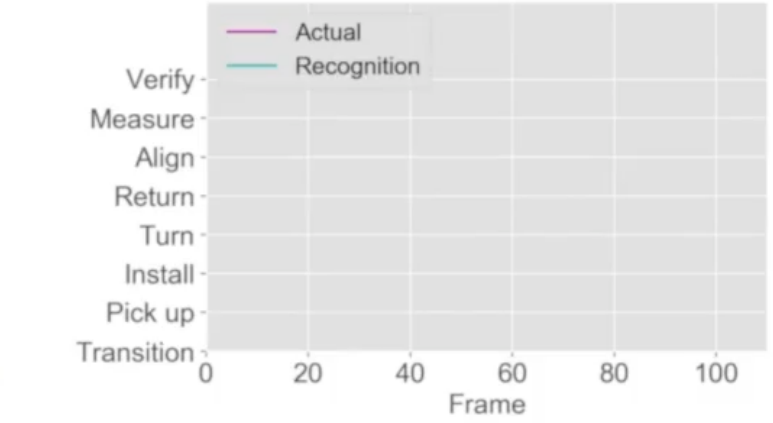
NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

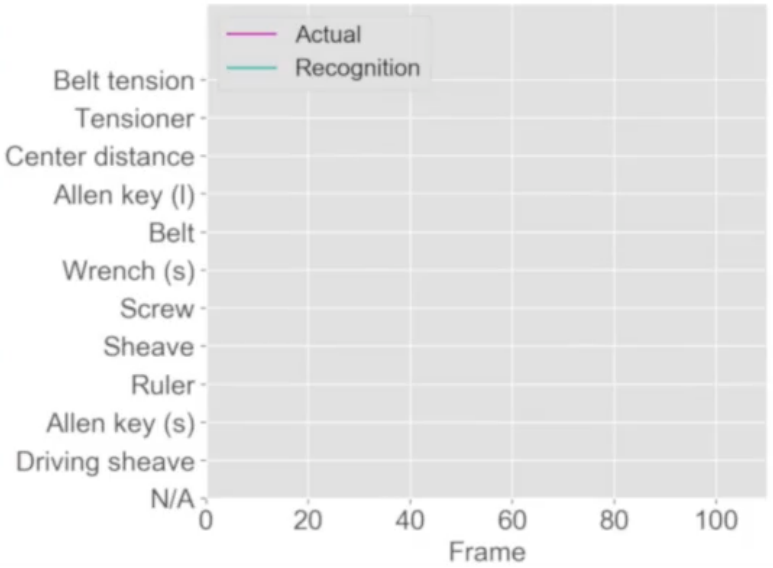
Experimental Evaluation: Action Prediction (2)

Action recognition



Object recognition





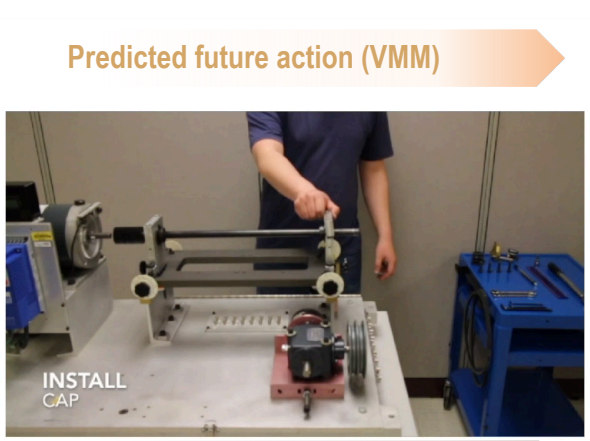
Overall recognition accuracy: 94.7%

Prediction time avg: <0.01s

NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Experimental Evaluation: Action Prediction (3)

- Future action prediction by VMM, based on current and past actions recognized by CNN



Predicted future action (VMM)

Current and past actions (DCNN)

Pick up screw | Future action

Install cap | Current action

Pick up cap

Install cap

Pick up cap

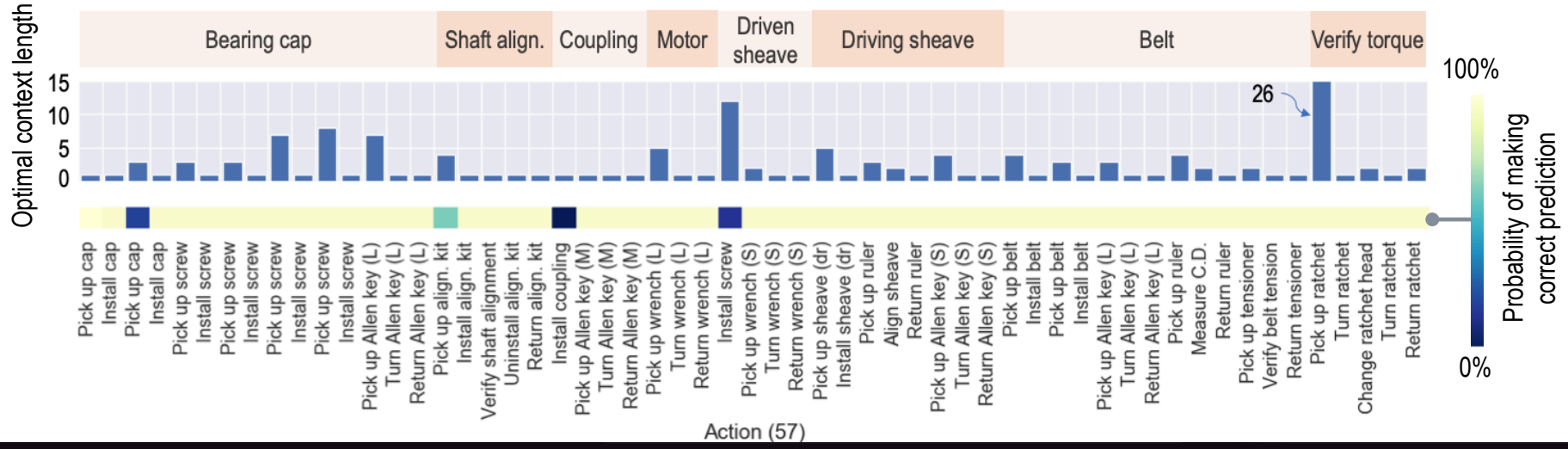
Start | Past action

From VMM

From CNN

Overall prediction accuracy: 97.5%
Prediction time avg: <0.01s

Indicates historical action used by VMM for prediction



NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

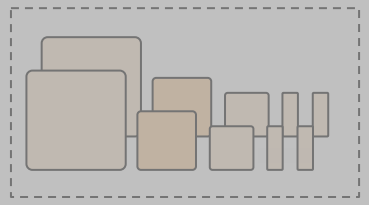
Sensing, Cognition, and Prediction

Sensing

Action Recognition

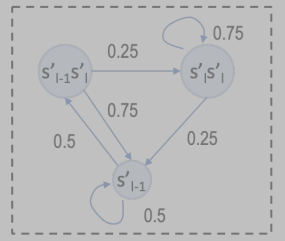
Action Prediction

Video recording



Deep Convolutional Neural Networks

Sequence of current and past actions (e.g. $s'=s'_1,s'_2,\ldots,s'_l$)

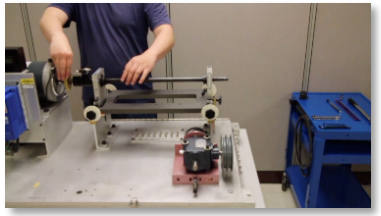


Variable-order Markov Model

Predicted future action

$\arg\max_w p(w|s'_{l-L+1},\ldots,s'_l)$

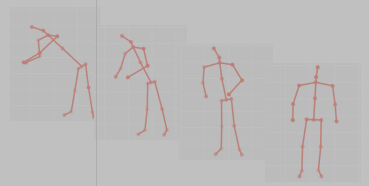
Pass *which* tool to worker?



Workspace

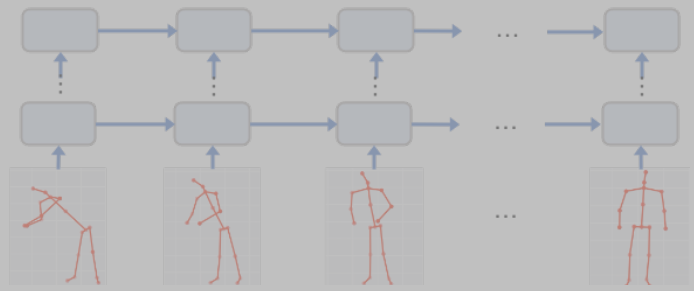
Action

Trajectory Prediction



Pose tracking

Historic trajectory of human pose



Recurrent Neural Networks

Predicted future human pose

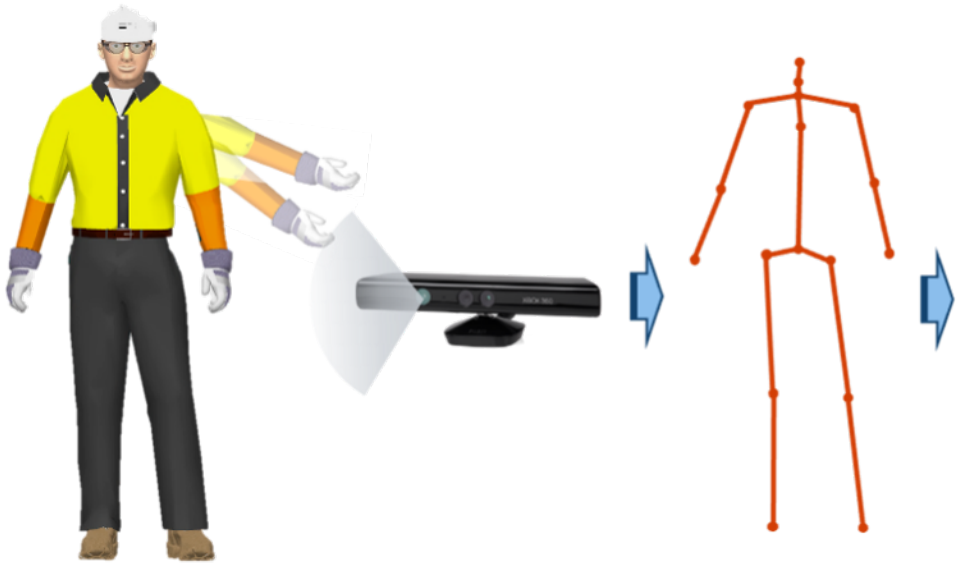
Where to pass the tool?

NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Trajectory Prediction

Recurrent Neural Network

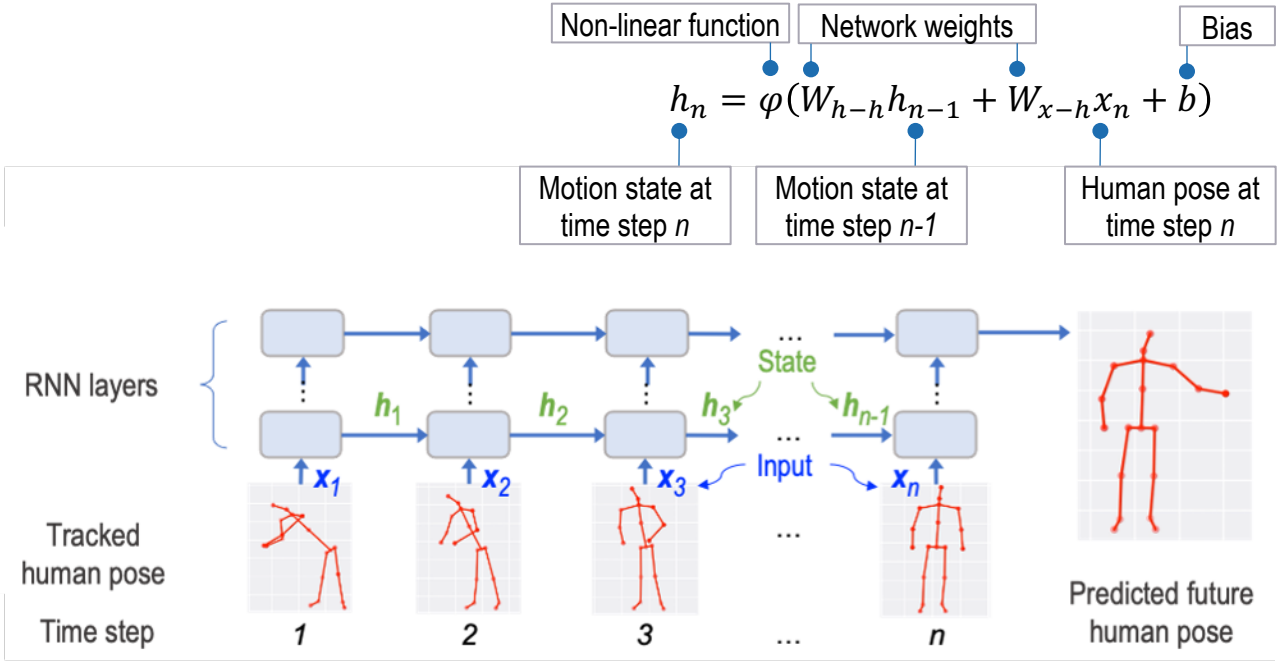
- ✓ Capture human motion evolution pattern by concurrently analyzing influence of motion state (extracted from coordinates of body joints) at each of time steps on state at subsequent step



Human worker
(Workspace)

Kinect sensor

Human pose
(Skeleton model)



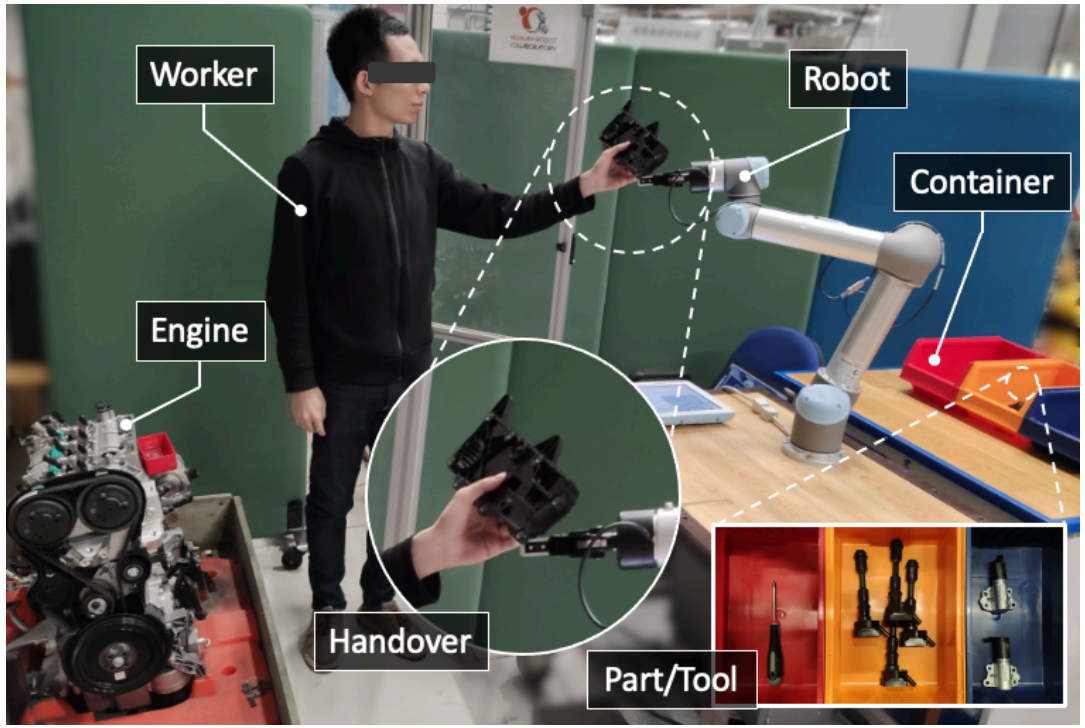
Trajectory prediction based on Recurrent Neural Network

NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

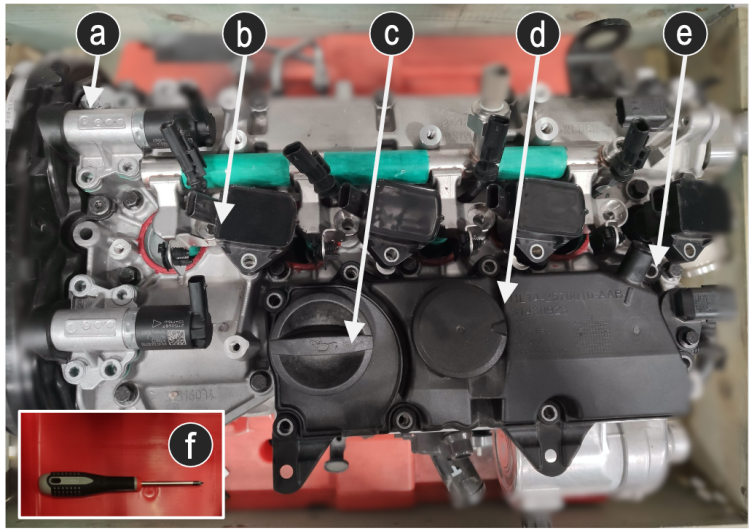
Experimental Evaluation: Trajectory Prediction (1)

Engine assembly:

- Parts/tools required: a large cover, a cover cap, two wire collectors, four plugs, a screwdriver, and eleven screws
- Human pose tracked by Kinect sensor



Human-robot collaborative workspace



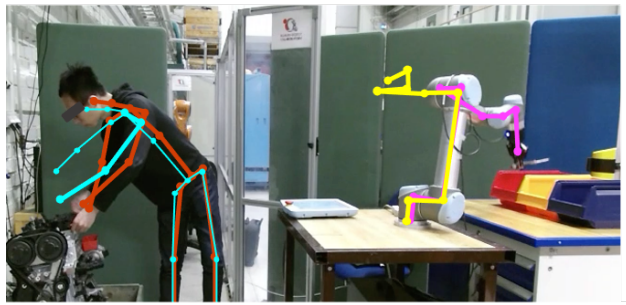
Engine assembly detail

NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

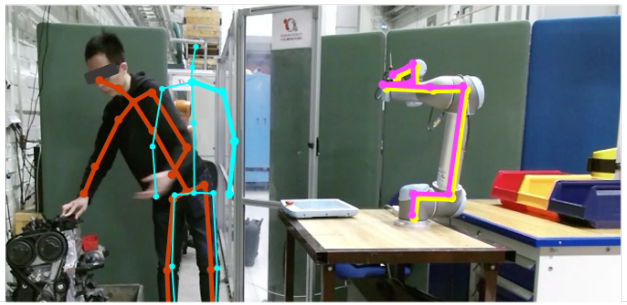
Experimental Evaluation: Trajectory Prediction (2)

Collaboration between human & robot presented by:

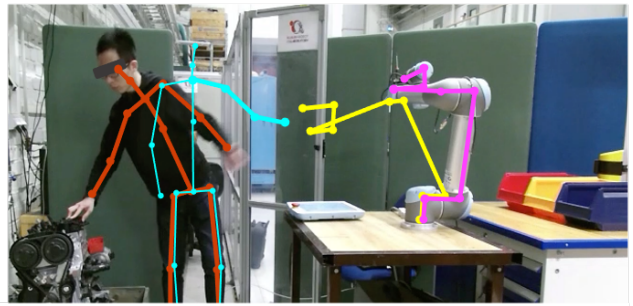
- Robot responding to predicted motion trajectory, following in real time human hand as it extends for handover
- Robot picking up relevant part/tool during installation



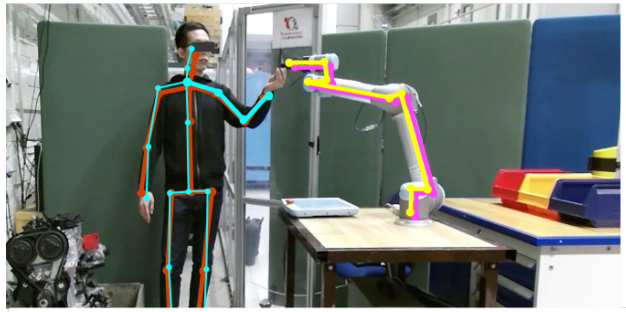
Robot detects worker as installing large cover, robot triggered to pick up screwdriver



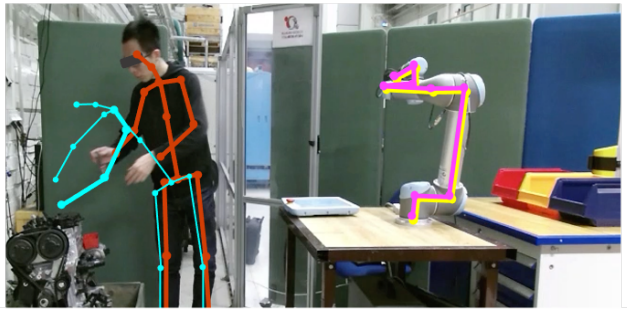
Worker completes installation, robot predicts worker moving to standing position



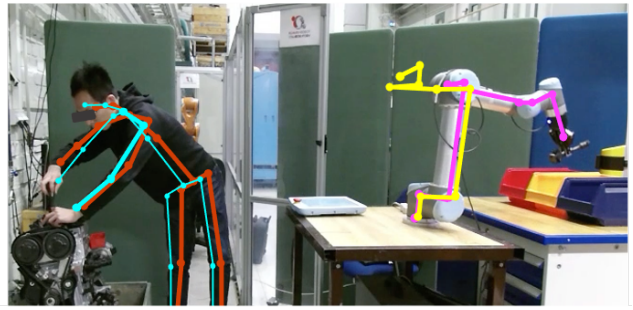
Robot predicts worker to move to handover, robot triggered and moves to predicted location



Robot adjusts handover location until worker's arm stops for handover



Worker returns to installation position and resumes engine assembly



Robot detects worker as installing, robot triggered to pick up next part

Actual human pose: red; Predicted human pose: cyan; Actual robot position: pink; Target robot position: yellow

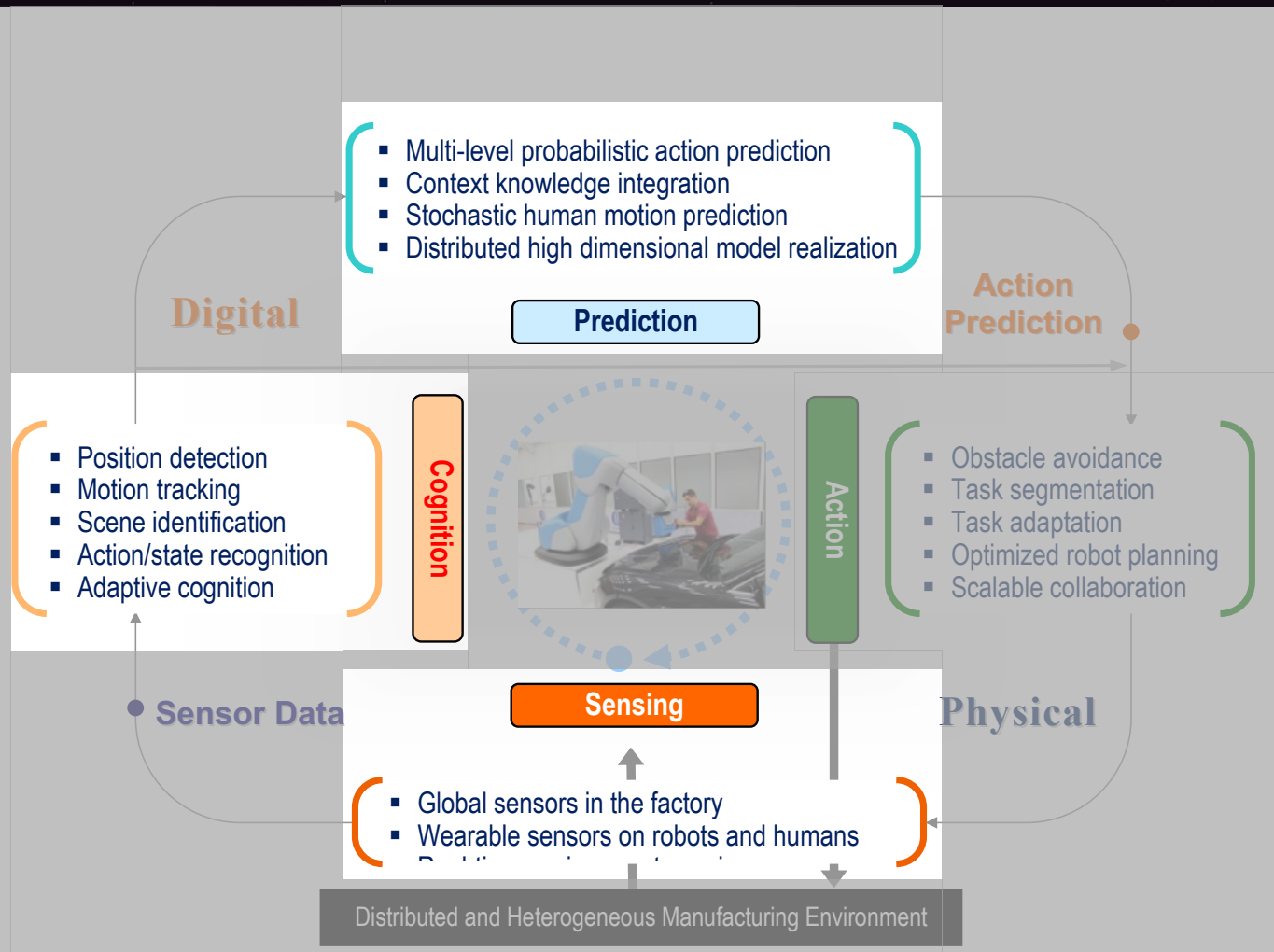
NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

NRI: INT: COLLAB: Manufacturing USA: Intelligent Human-Robot Collaboration for Smart Factory

Research Goals:

- We propose four research components to realize the envisioned human-robot collaboration (HRC) for an automated HRC manufacturing cell,
  - from data acquisition in the physical domain
  - to data manipulation in the digital domain, and
  - back to robot control in the physical domain:

sensing, cognition, prediction & action.



NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Action

Task Objective:

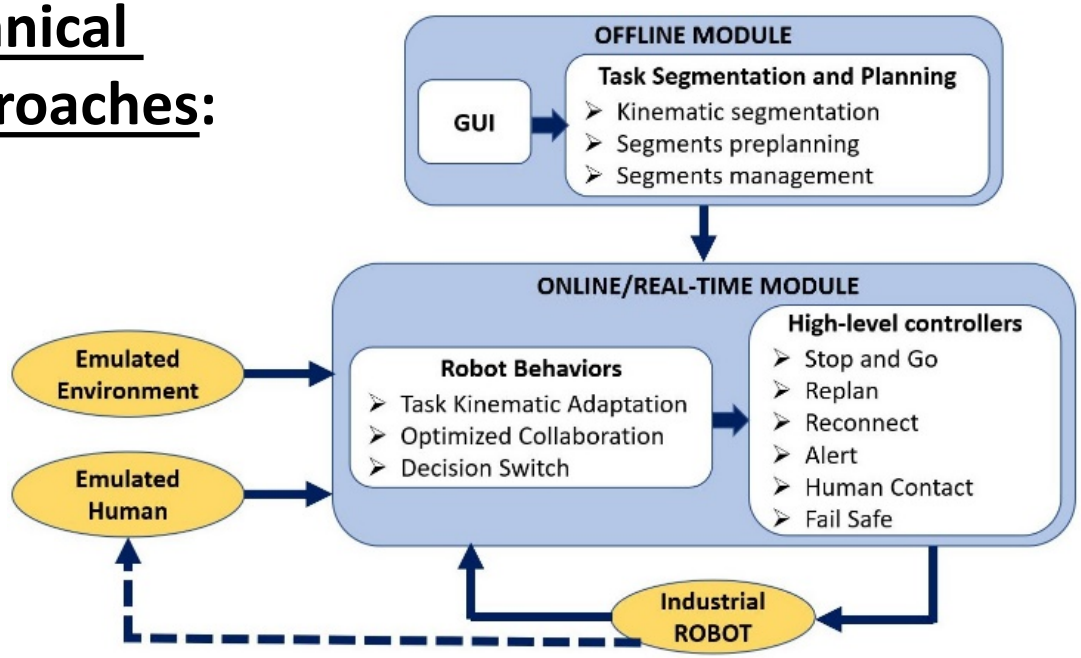
Seamless integration of sensing, cognition and prediction into the **robot controller**. To provide the robot **Proactive Adaptive Collaboration Intelligence (PACI)** and switching logic within its control architecture in order to give the **robot the ability to optimally and dynamically adapt its motions**, given a priori knowledge and predefined execution plans for its assigned tasks.

Challenge: **Augmenting** the robot's **decision-making process** to have greater situation awareness and **to yield smart robot behaviors/reactions** when subject to different levels of human-robot interaction, **while maintaining safety and production efficiency**.

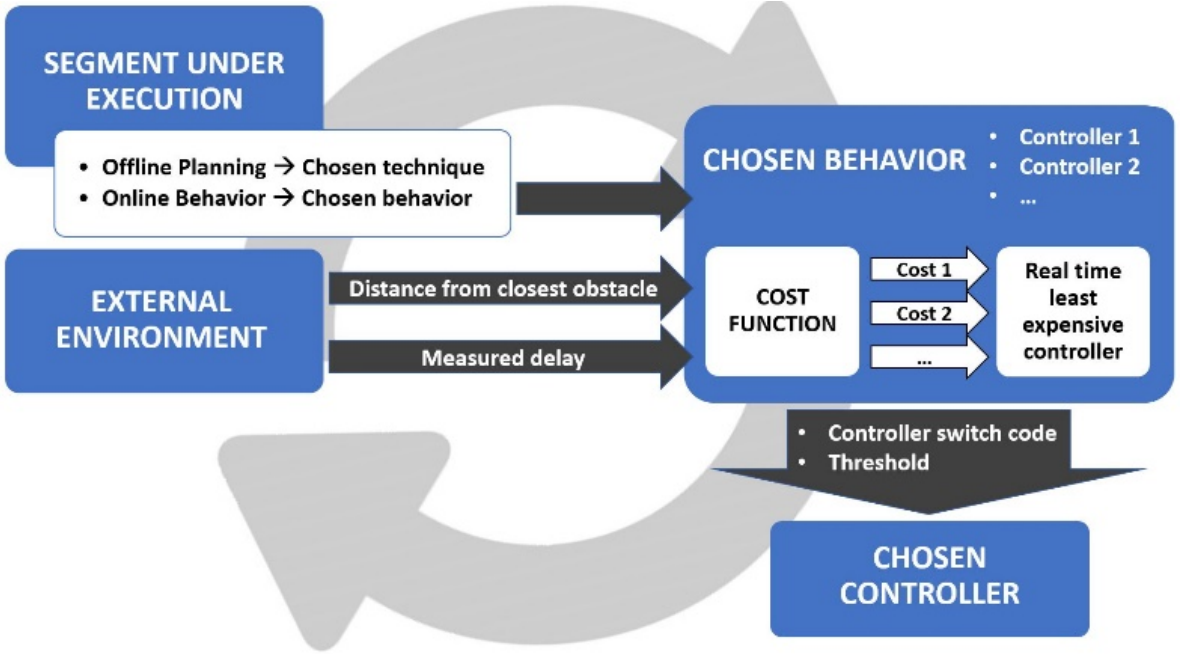
NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Action

Technical Approaches:



Control architecture.



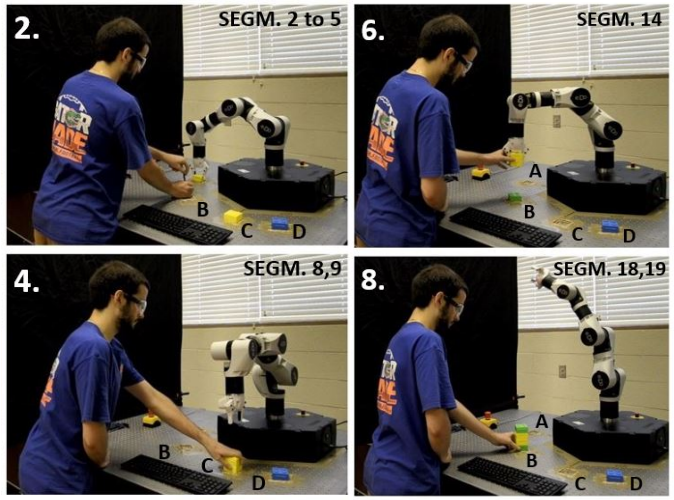
Working principle of online switching logic.

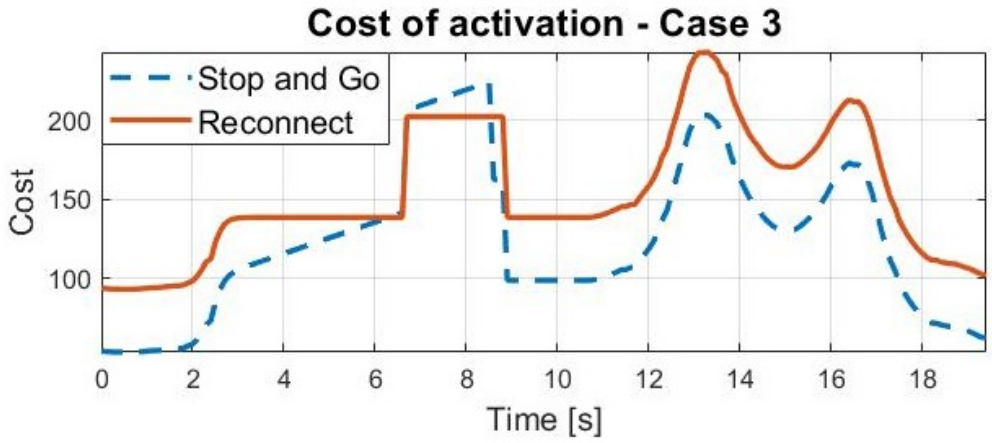
Task Segmentation and Planning: Offline module is equipped with a GUI that takes as input the requests of the user and feeds the processed information to a second module, which is in charge of the kinematic segmentation of the task(s) and the preplanning and management of the created segments.

NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Action

Example:





Cost trends – trajectory replan due to human presence detected at t=6.8 seconds.

Reactive Behaviors: Online/Real-time module kinematically adapts robot motion in accordance to preplanned segmentation behavior priorities, and predicted and sensed human motion and environment. Robot reactive behaviors are achieved via cost function-based switching logic activating the best suited high-level controller (*stop-n-go, replan, reconnect replan, alert, human contact*). The PACI's underlying segmentation and switching logic framework yields a high degree of modularity and flexibility.

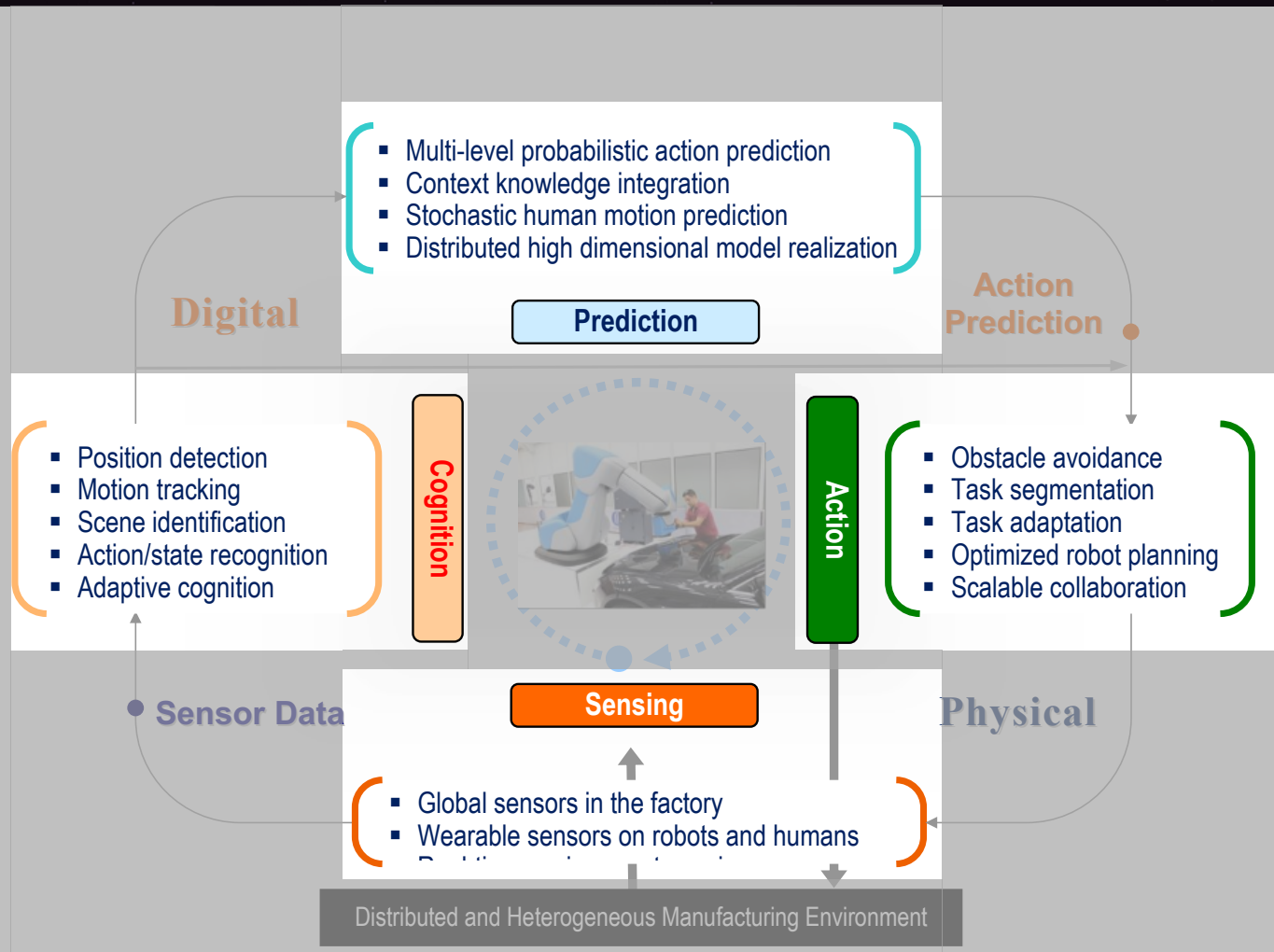
NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

NRI: INT: COLLAB: Manufacturing USA: Intelligent Human-Robot Collaboration for Smart Factory

Research Goals:

- We propose four research components to realize the envisioned human-robot collaboration (HRC) for an automated HRC manufacturing cell,
  - from data acquisition in the physical domain
  - to data manipulation in the digital domain, and
  - back to robot control in the physical domain:

sensing, cognition, prediction & action.



NSF/NRI 2.0: ENG/CMMI and CISE/IIS
Award ID#: 1954548/1830383/1830295

Thank you!

Welcome to our poster for questions & comments!