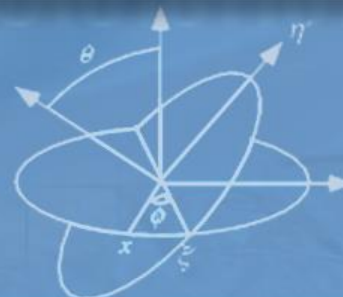# A Detection-based Approach to Multiview Action Classification in Infants

Carolina Pacheco[1], Effrosyni Mavroudi[1], Elena Kokkoni[2], Hebert G. Tanner[3] and René Vidal[1]

[1] Johns Hopkins University, [2]University of California Riverside, [3]University of Delaware

JHU vision lab

JOHNS HOPKINS
MATHEMATICAL INSTITUTE
for DATA SCIENCE

THE DEPARTMENT OF BIOMEDICAL ENGINEERING
The Whitaker Institute at Johns Hopkins

Center for
IMAGING
SCIENCE

Motivation

Activity recognition in children and infants

Safety monitoring

Source: Fang et al.

Object-play behavior assessment
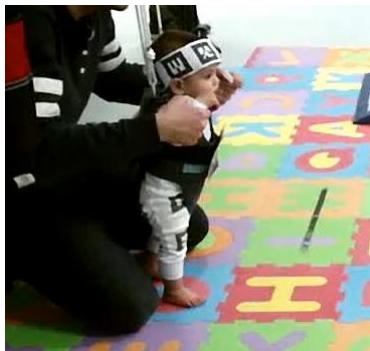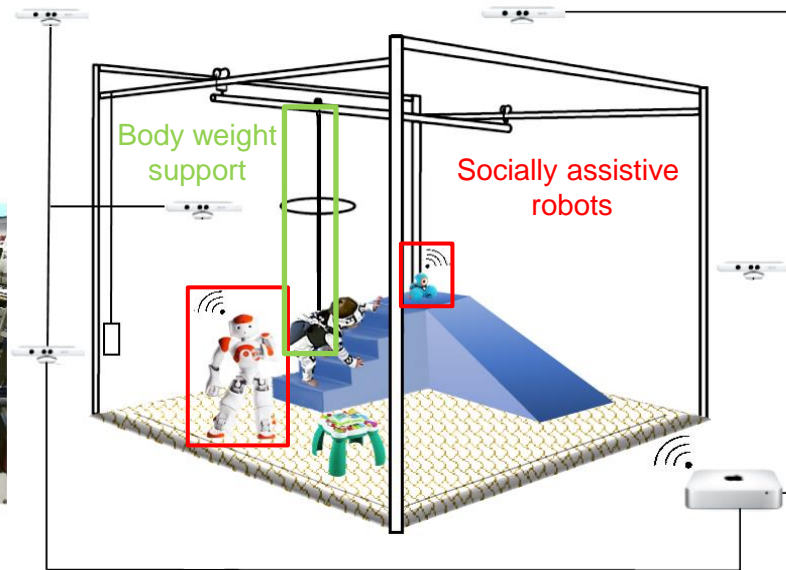

Source: Westeyn et al.

Child-robot interaction

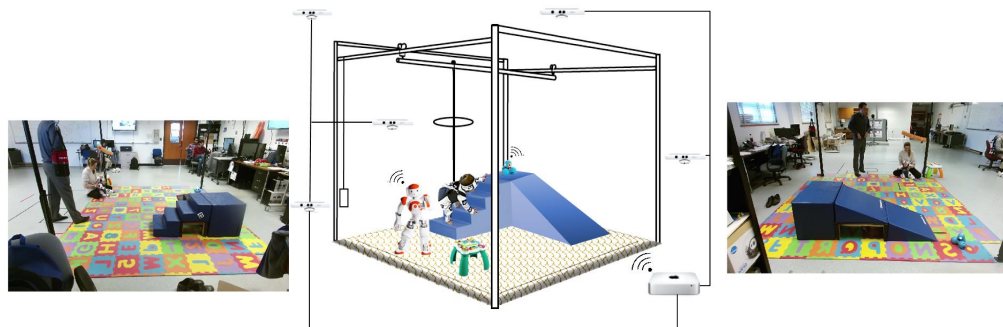Activity recognition in infants is different from activity recognition in adults

Source: https://youtu.be/GBkJY86tZRE

[1] Fang et al., "A vision-based infant respiratory frequency detection system"
[2] Westeyn et al., "Monitoring children's developmental progress using augmented toys and activity recognition"
[3] Suzuki et al., "Activity recognition for children using self-organizing map"
[4] Efthymiou et al.,"Multi-view fusion for action recognition in child-robot interaction"

JOHNS HOPKINS
MATHEMATICAL INSTITUTE
for DATA SCIENCE

Body weight support

Socially assistive robots

[5] Kokkoni et al., "Gearing smart environments for pediatric motor rehabilitation"

JOHNS HOPKINS
MATHEMATICAL INSTITUTE
for DATA SCIENCE

Challenges and Contributions

Pediatric learning environment
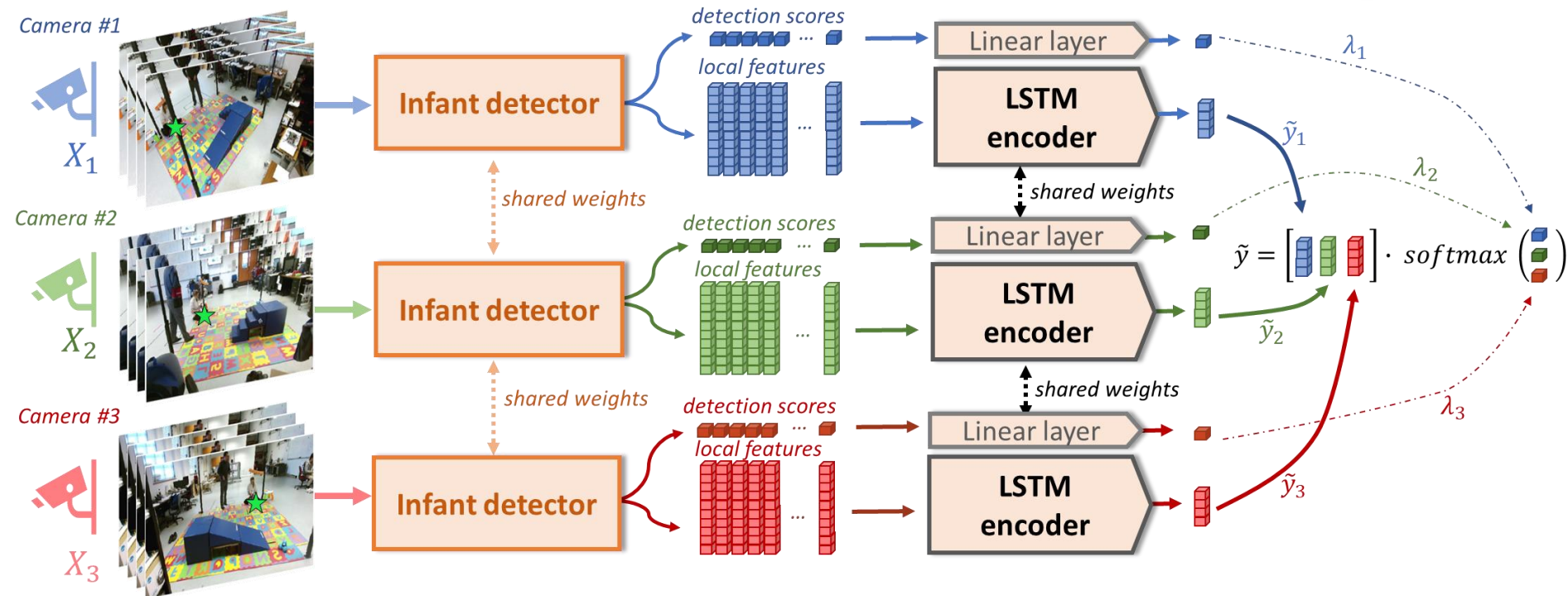

Challenges

* Off-the-shelf pre-trained systems wouldn't work
* Infants are often occluded by other elements or actors
* Infants' body size is small, covering in average <1.7% of frame area.

Contributions

✓ A detection-based multiview action classification system
✓ A novel and meaningful use of the detection confidence scores
✓ State-of-the-art action classification results for rehabilitation therapy in infants

[5] Kokkoni et al., "Gearing smart environments for pediatric motor rehabilitation"

JOHNS HOPKINS
MATHEMATICAL INSTITUTE
for DATA SCIENCE

Detection-based Multiview Action Classification

✓ Architecture is view-independent
✓ Features are close to view-invariant
✓ Model can handle new or missing views at test time
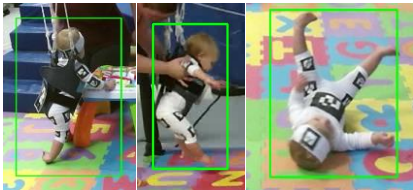
Experiments

Data acquisition

6 infants (7.8 - 23.7 months-old)
8 1-hour sessions per subject

Annotation
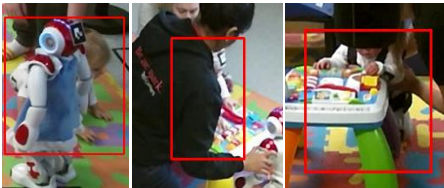
Spatial localization
Actions

Spatial localization

High visibility

Medium visibility

Low visibility

Actions

Crawling

Sitting

Standing

Walking

JOHNS HOPKINS
MATHEMATICAL INSTITUTE
for DATA SCIENCE

Experiments

Data acquisition

6 infants (7.8 - 23.7 months-old)
8 1-hour sessions per subject

Annotation
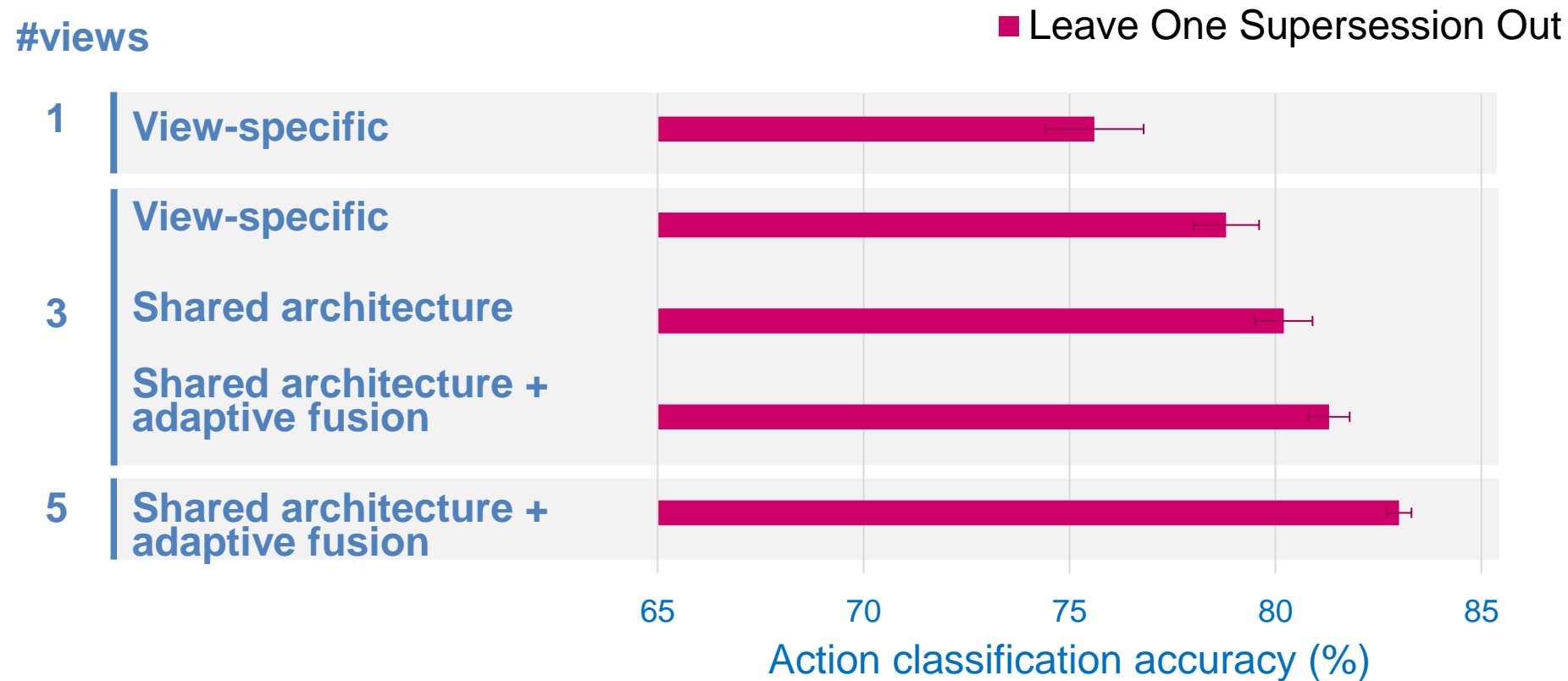
Spatial localization
Actions

Training

Step 1. Fine-tune view-agnostic Faster-RCNN detector
Step 2. Train multiview action classifier (LSTM)

Cross-validation
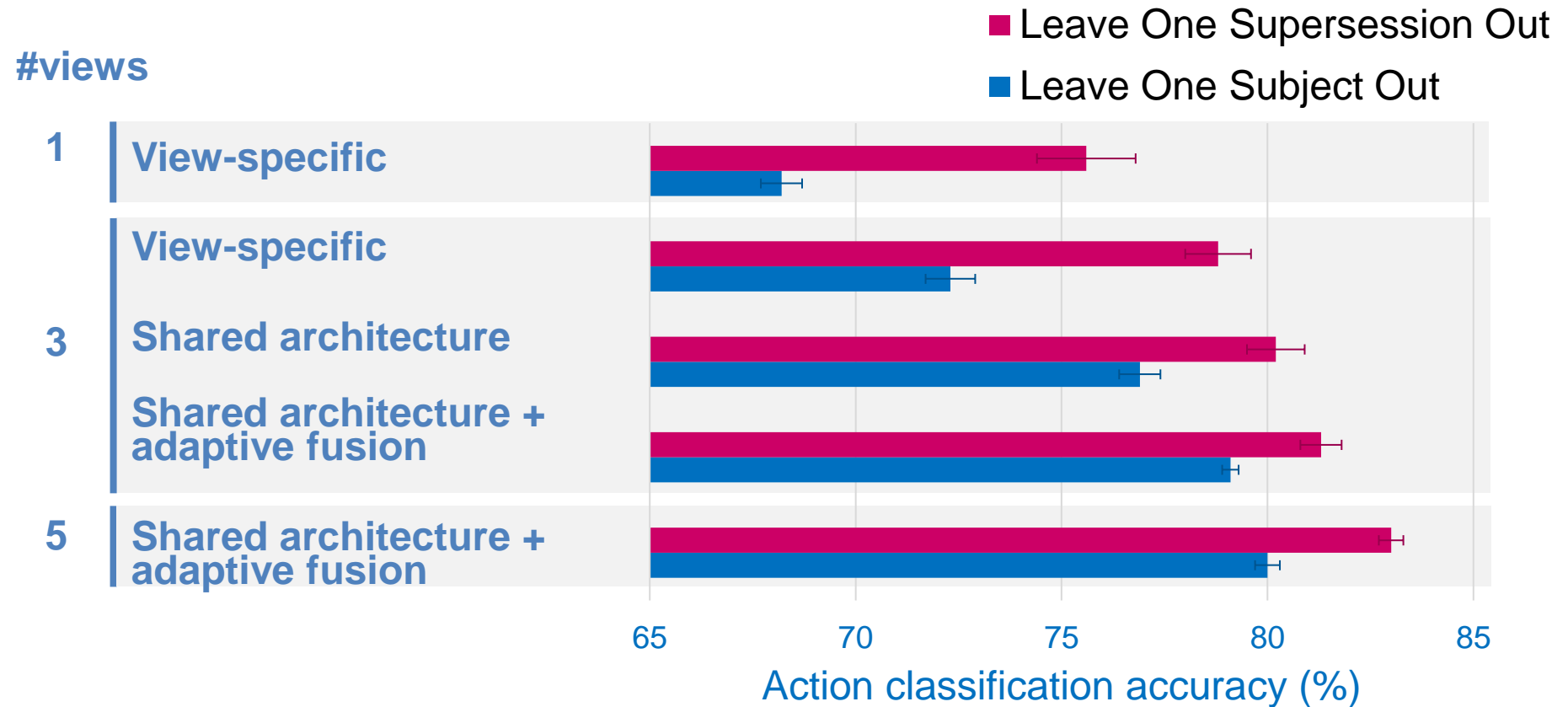
Leave one super-session out
Leave one subject out

Main results

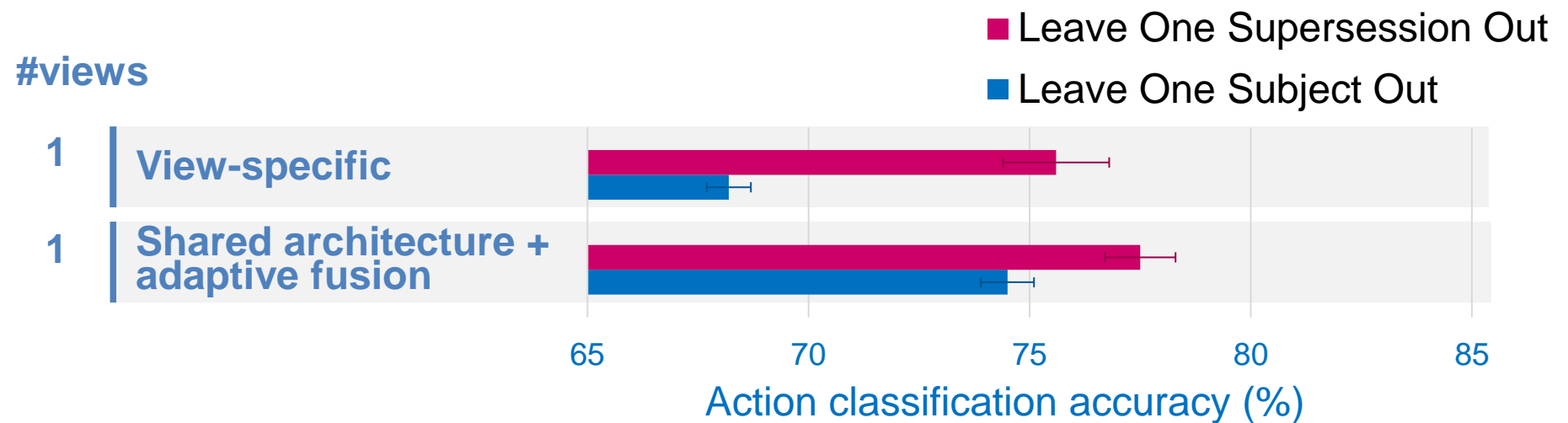The proposed method **outperforms all other multiview strategies** studied

Main results

Main results

The multiview system **increases the performance** even when a **single camera** is available at test time

#views

Legend:
- ■ Leave One Supersession Out
- ■ Leave One Subject Out

#views		Action classification accuracy (%)
1	View-specific	
1	Shared architecture + adaptive fusion	

X-axis: Action classification accuracy (%) — 65, 70, 75, 80, 85

JOHNS HOPKINS
MATHEMATICAL INSTITUTE
for DATA SCIENCE

Main results

The proposed system **significantly outperforms** prior work on action classification for rehabilitation therapy in infants

[5] Kokkoni et al., "Gearing smart environments for pediatric motor rehabilitation"
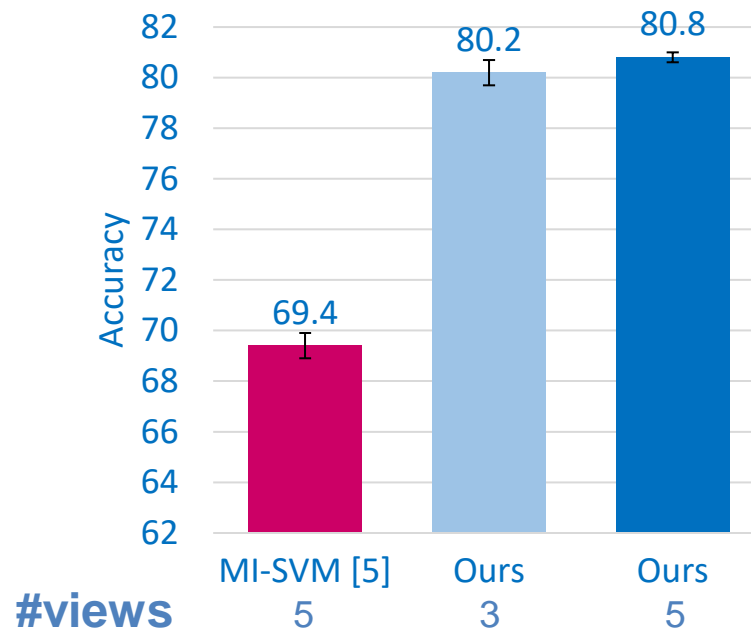
A Detection-based Approach to Multiview Action Classification in Infants

Thank you!

Paper #1669
Poster Session 2.1

Contact: cpachec2@jhu.edu

JOHNS HOPKINS
MATHEMATICAL INSTITUTE
for DATA SCIENCE

Center for
IMAGING
SCIENCE