BOSTON UNIVERSITY

Action Detection for Behavioral Analysis on RABC Dataset

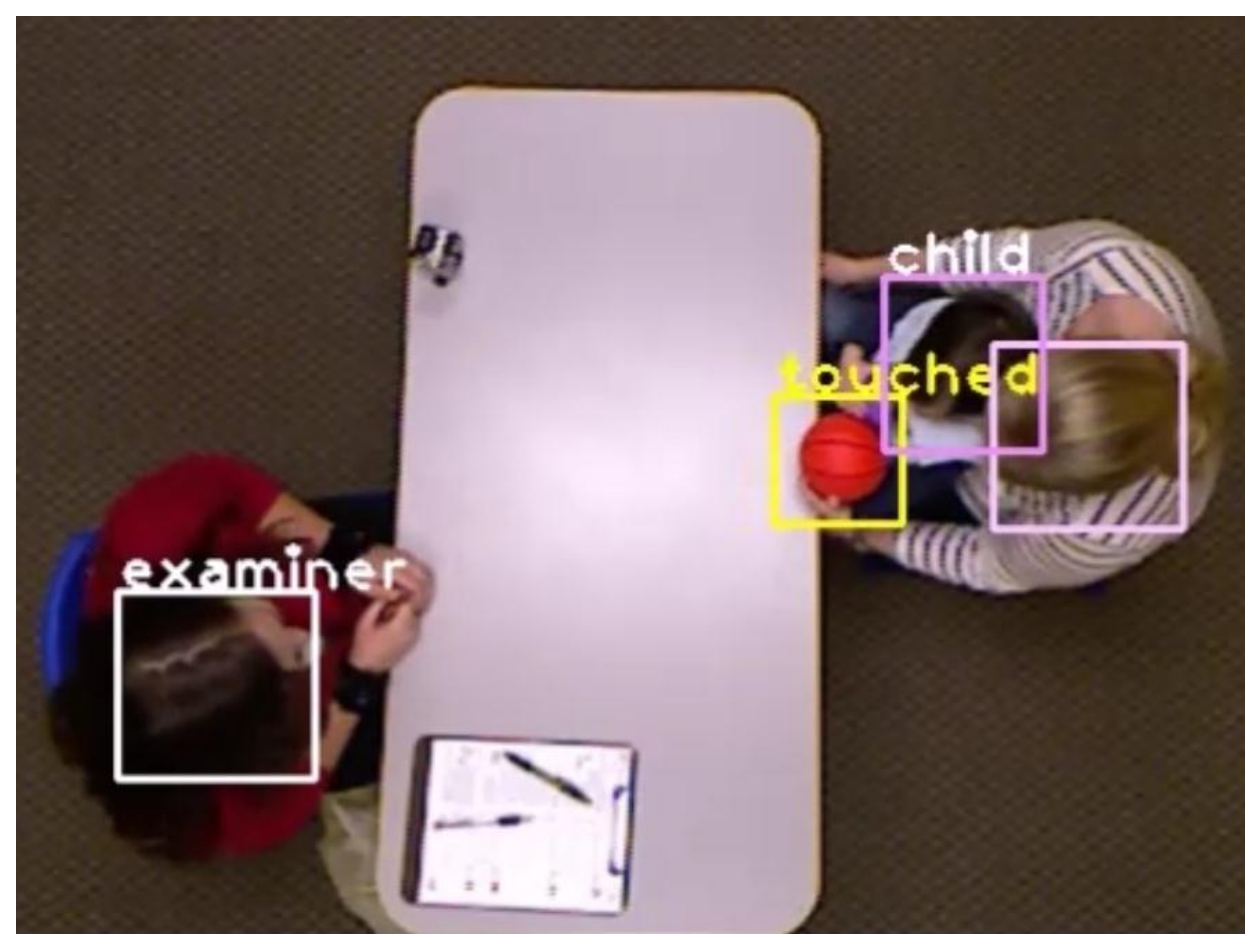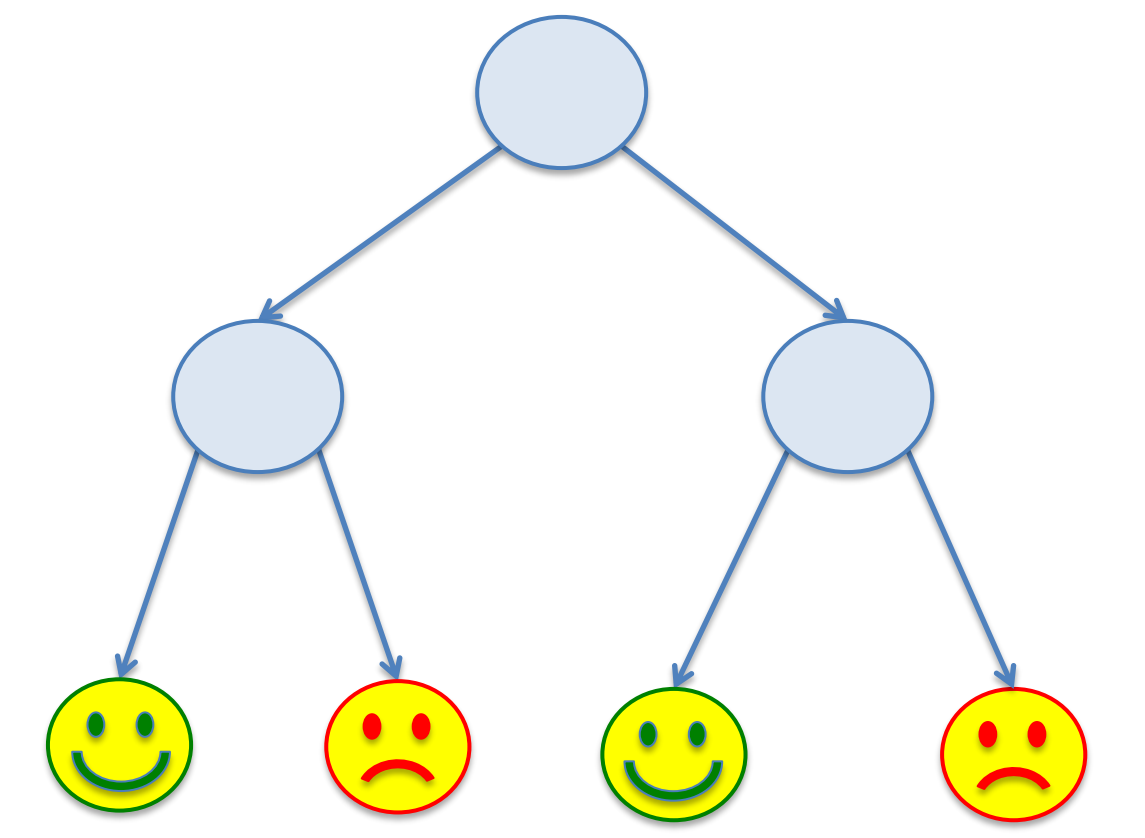Jianming Zhang   Liliana Lo Presti   Stan Sclaroff

Action & Event Detection Were Shown to Be Useful in Behavioral Analysis

Detecting & tracking objects

Features

Ball visibility	# frames the ball is detected
Ball touched	# frames the ball is touched
Ball touched by child	# frames the child touches…
…..	……

Feature extraction based on tracking and action detection results
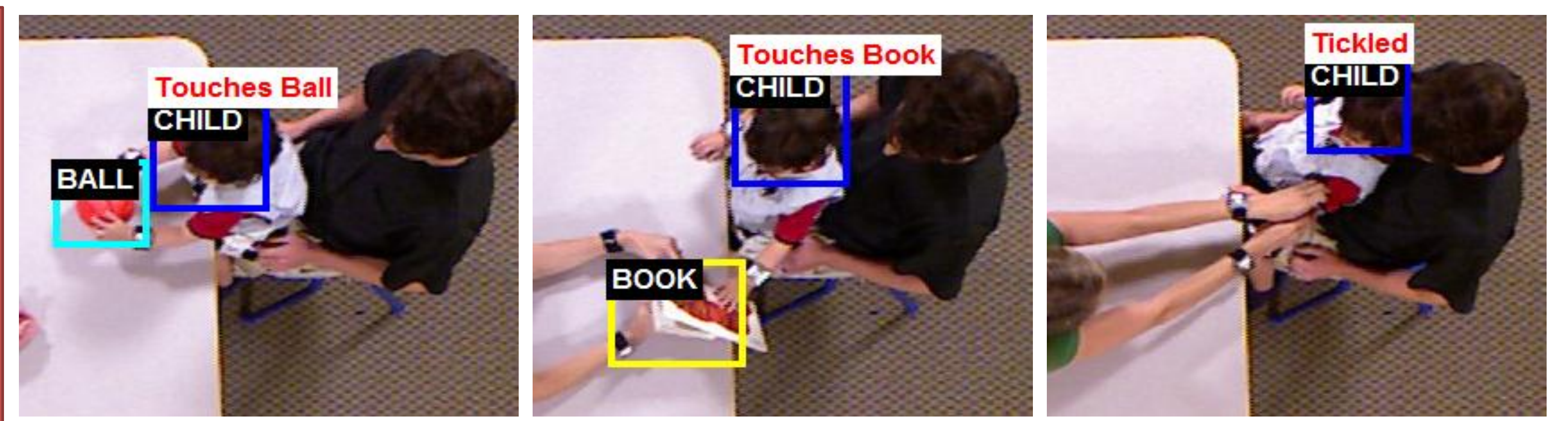
Easy to engage vs. not easy to engage

Using features extracted based on the basic action and event detection results leads to some interesting engagement prediction results [2].

We Provide More Action Detection Results for RABC Dataset

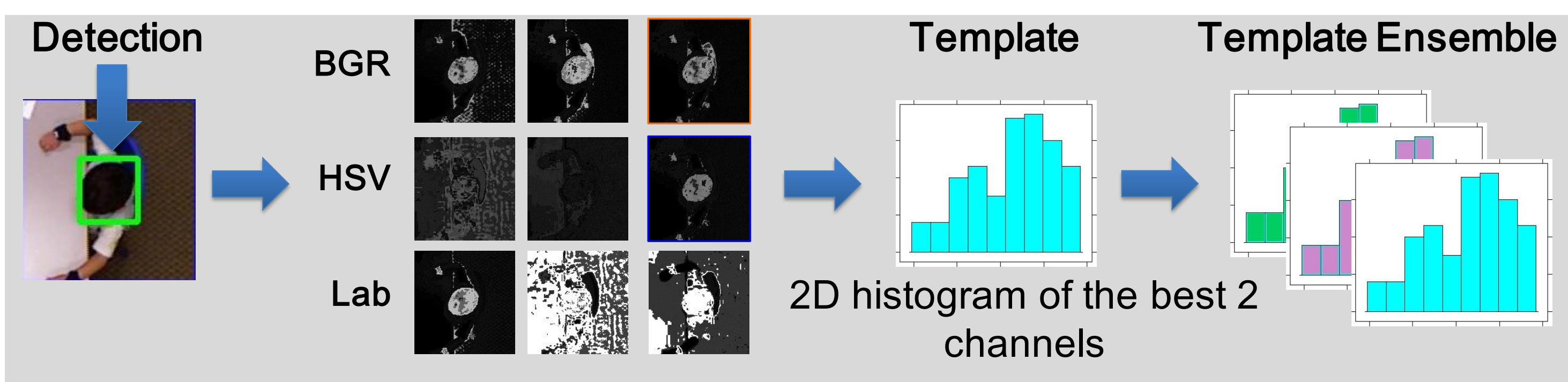■ **[Reaching/Pointing/Pushing] Actions**

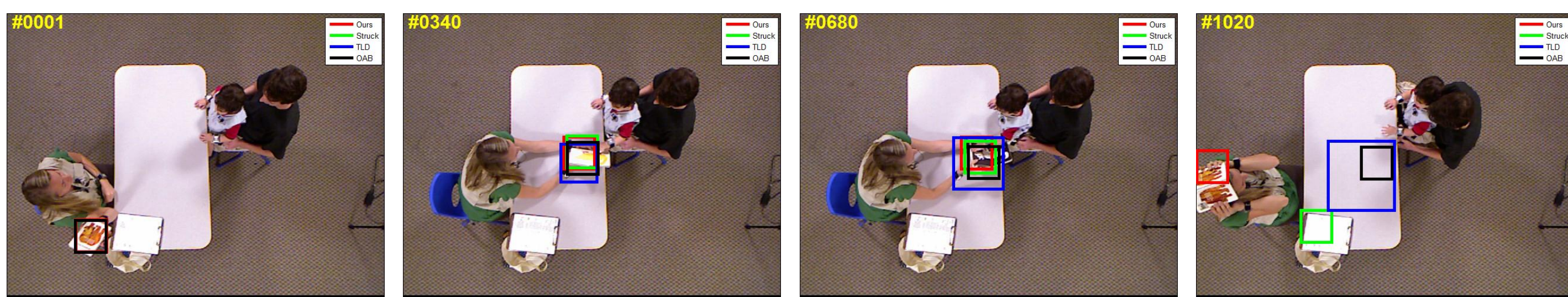■ **Ball_touch/Book_touch/Tickle Actions**

Technical Approach

■ Improved automatic object detection & tracking software for RABC Kinect data based on [1]. (shared with Gatech for gaze estimation.)

Detection   BGR   HSV   Lab   Template   Template Ensemble

2D histogram of the best 2 channels

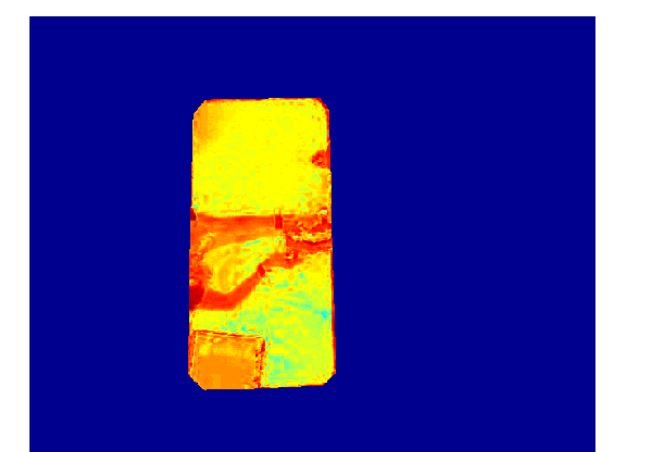■ A new tracker that is base on expert ensemble and minimum entropy criterion.

$$\hat{E}_t = \arg\min_{E \in \mathbf{E}} \sum_{k \in [t-\Delta, t]} \mathcal{L}_E^k,$$

$$\mathcal{L}_E(\mathbf{x}, \mathbf{z}) = -L(\theta_E; \mathbf{x}, \mathbf{z}) + \lambda H(\mathbf{y}|\mathbf{x}, \mathbf{z}; \theta_E)$$

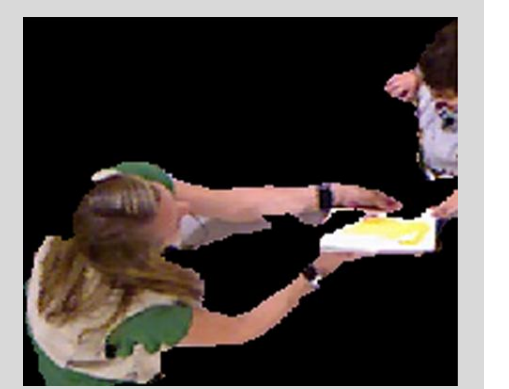■ Body segmentation: A preliminary step for many other techniques;

-- Segment the body parts above the table using depth data;
-- Detect salient regions on the table using color whitening [3];

$$\text{Saliency} = \left\| [(r, g, b) - (\bar{r}, \bar{g}, \bar{b})] \cdot \text{Cov}_{rgb}^{-1} \right\|_2$$
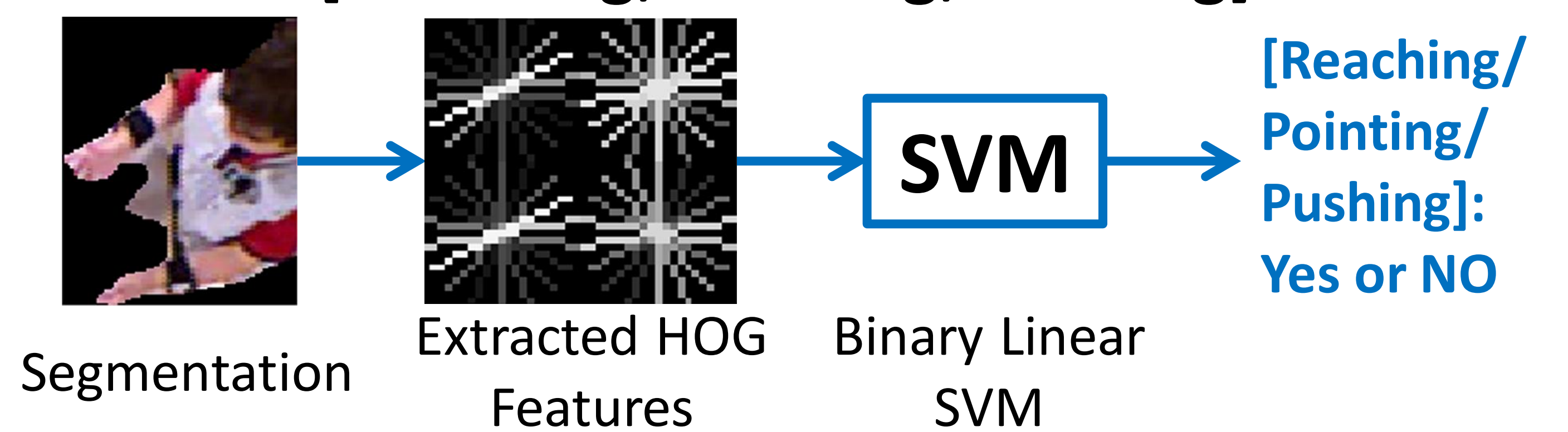
-- Combine the above results to produce connected regions of the child's upper body.

Example: the connectivity of the child's body to some object can indicate some interaction (book_touch).

■ Detect [Reaching/Pointing/Pushing] actions

Segmentation   Extracted HOG Features   Binary Linear SVM

SVM   **[Reaching/ Pointing/ Pushing]: Yes or NO**

Future Work

■ To leverage depth, color and body structural information for robust hand tracking will be one future direction, because hand trajectories contain important cues for behavioral analysis.

■ To provide more and better action detection results of wide interests.

■ To build behavioral primitives by combining our and other groups' results.

[1] J. Zhang, L. Lo Presti, S. Sclaroff, "Online Multi-Person Tracking by Tracker Hierarchy", In Proc. IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS), 2012.

[2] J. M. Rehg, et. al., "Decoding Children's Social Behavior", In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013.

[3] A. Garcia-Diaz, et. al., "Saliency from hierarchical adaptation through decorrelation and variance normalization" Image and Vision Computing, 2012: 51-64.