

# Activity Detection in Extended Video using Action Tubelets

Chihiro Shiraishi<sup>1</sup>, Nakamasa Inoue<sup>1</sup>, Aleksandr Drozd<sup>1</sup>,  
Koichi Shinoda<sup>1</sup>, Shi-wook Lee<sup>2</sup>, and Alex Chichung Kot<sup>3</sup>

Tokyo Institute of Technology<sup>1</sup>,  
National Institute of Advanced Industrial Science and Technology<sup>2</sup>,  
Nanyang Technology University<sup>3</sup>

We propose a system for activity detection which utilizes the Action Tubelet Detector (ACT-detector).  
By introducing a threshold we can reduce the false alarm rate for videos with sparsely distributed activities.

## Task: Activity Detection

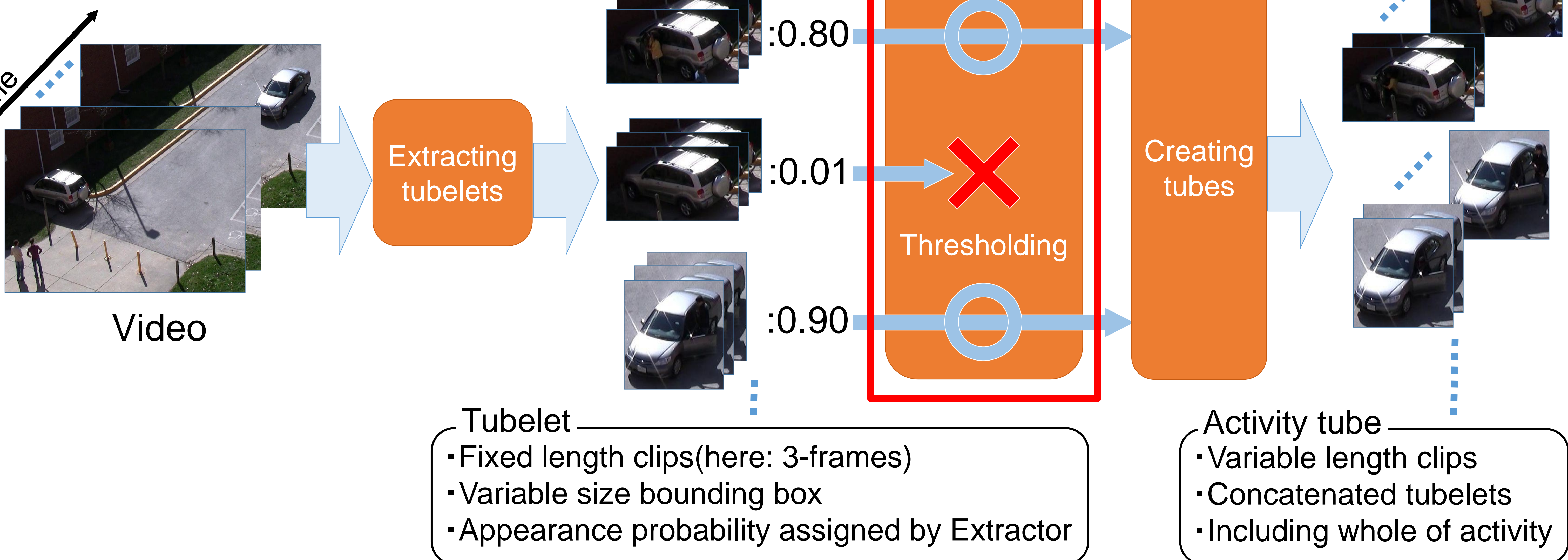
This task handles 19 target activities related to “Person” and “Vehicle”.  
Examples: Talking, Opening Trunk, Loading of a Vehicle,...  
The system automatically detects and temporally localizes all activities.

**Activity** —  
One or more people performing a specified movement or interacting with an object or group of objects.

## Method: Activity Tube Detector with Thresholding

Target activity:

A person exiting a vehicle

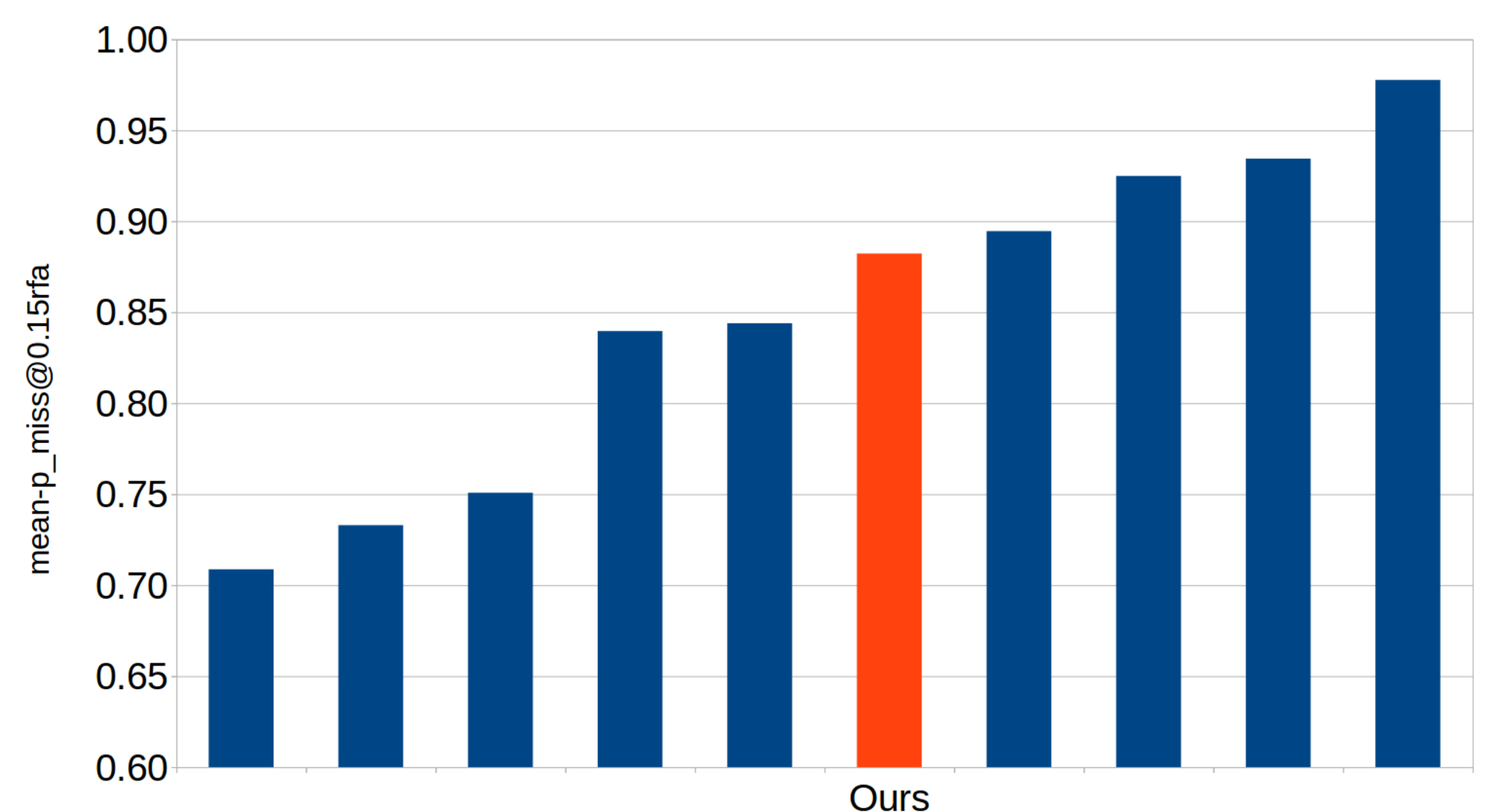


## Results

**mean-p\_miss@0.15rfa** —  
The mean probability of missed detection with a fixed rate of false alarm of 0.15 (1/min).

Comparing to the baseline (original ACT-detector)

Method	Dataset	
	Validation	Test
w/o thresholding (baseline)	0.885	—
w/ thresholding	<b>0.856</b>	0.882



Results of the AD Leaderboard Evaluation

## Conclusion & Future Work

We participated in the Activities in Extended Video (ActEV) challenge and evaluated our system using ACT-detector. Future work will focus on the detection of the beginning and ending of activities.