北京郵電大學
BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

BUPT-MCPRL@TRECVID 2015: Surveillance Event Detection

Zhicheng Zhao

Qi Chen, Xiang Li, Menglai Wang, Yanyun Zhao

BUPT-MCPRL

Beijing University of Posts and Telecommunications

Summary

- BUPT-MCPRL 2015 Retrospective Results

	Event	Our ADCR	Other Best Systems
CNN-based	Embrace	**0.7909**	0.8680
	ObjectPut	**1.0120**	1.0160
	Pointing	**1.0040**	1.0140
	PersonRuns	0.9700	**0.5768**
Trajectory-based	PeopleSplitUp	0.9387	**0.8934**
	PeopleMeet	1.0426	**0.8939**

Outline

- Retrospective System Overview
- Pedestrian Detection
- Detected by CNN
  - Embrace and Pointing
  - ObjectPut and PersonRuns
- Pedestrian Tracking
- Detected by Trajectory Analysis
  - PeopleMeet and PeopleSplitUp
- Conclusion

Retrospective System Overview

Pedestrian Detection by CNN

Embrace, ObjectPut, PersonRuns, Pointing

Classified by CNN → Events Fusion

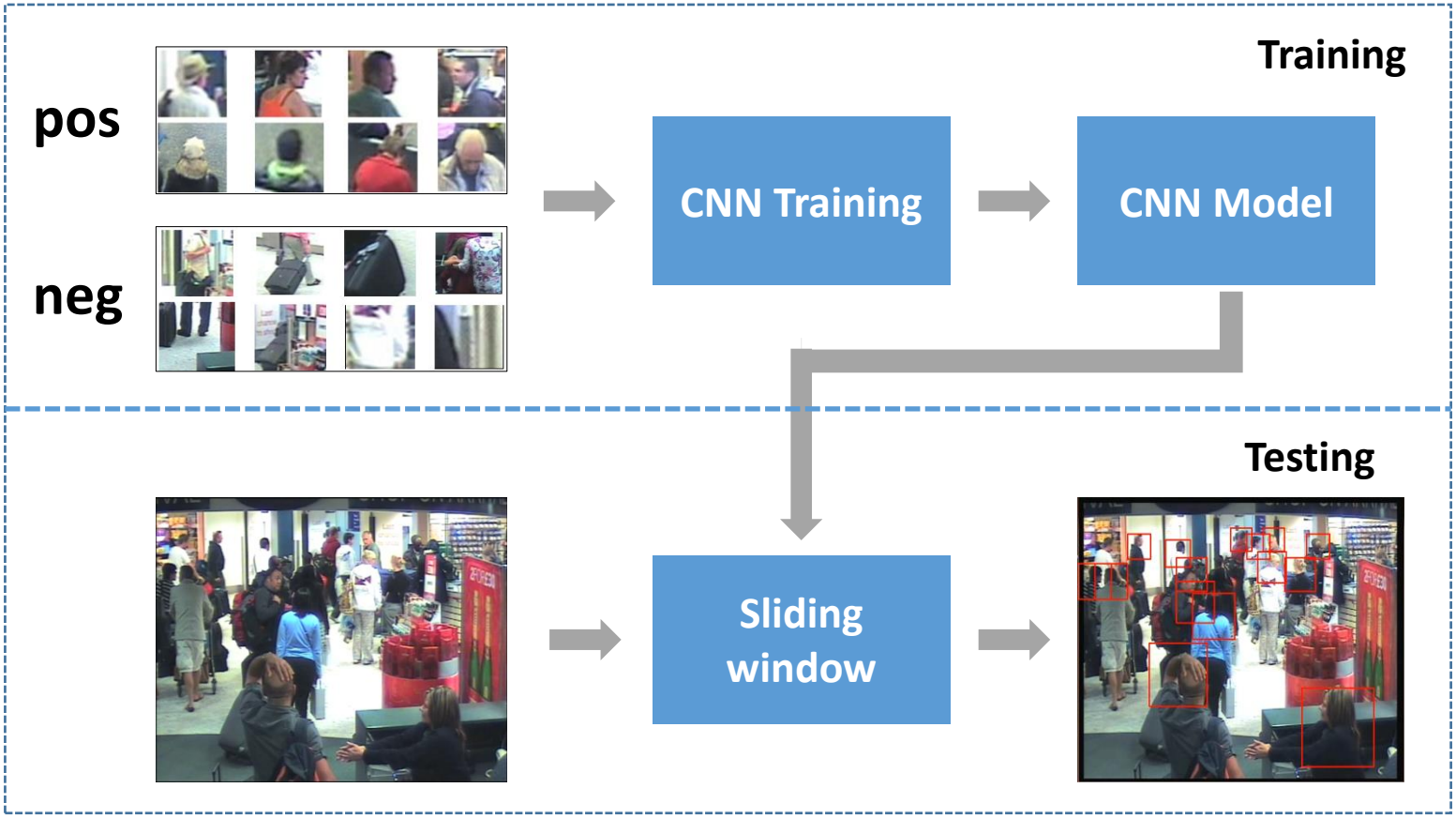
PeopleMeet, PeopleSplitUp

Pedestrian Tracking → Trajectory Analysis

Results

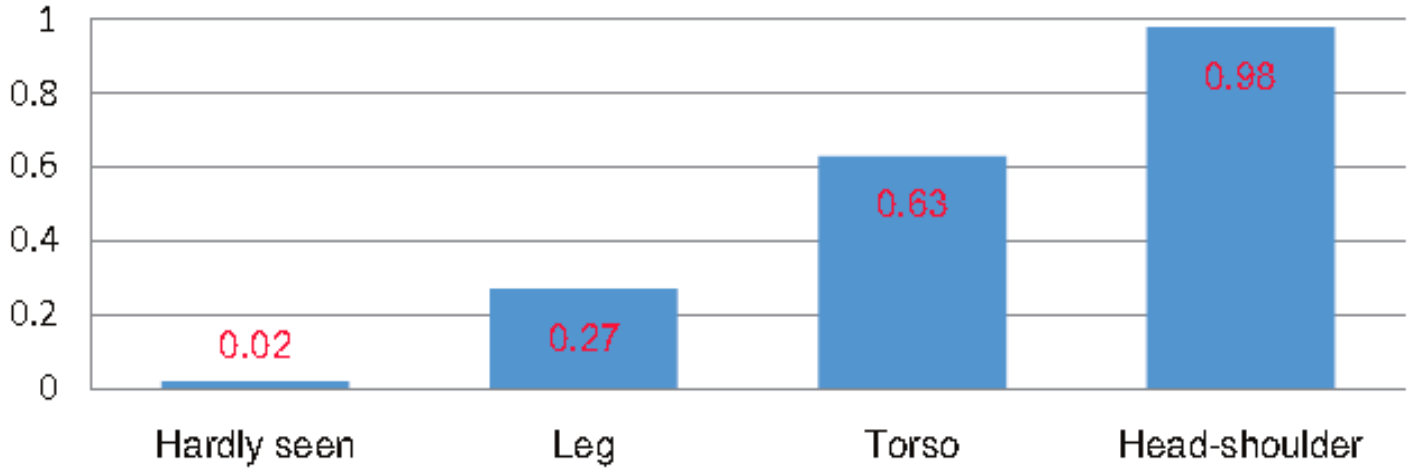
Pedestrian Detection

- Pedestrian Detection by a **H**ead-**S**houlder-CNN(**HsNet**)
  - Suppress the effect of partial occlusion



Pedestrian Detection

- **Why is head-shoulder?**
  - Most pedestrian instances (about 73%) show incomplete body parts
  - More than 98% keep head and shoulder
  - We detect the head-shoulder part instead of full-body for pedestrian detection



Pedestrian Detection

- **Why is head-shoulder?**

(Miss rate)

COMPARISON OF MR BY MODELING DIFFERENT BODY PARTS

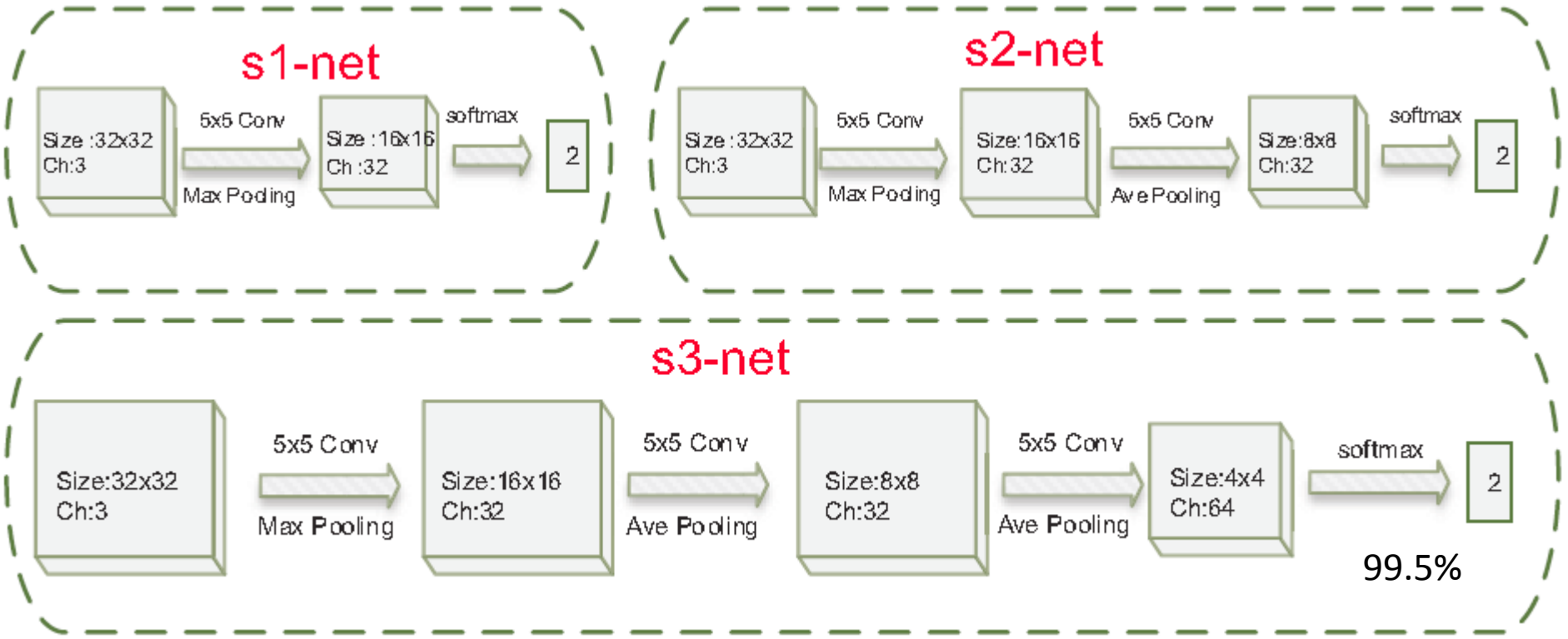
Model	Head-shoulder	Upper-body	Full-body
Cifar-10 Network	**54.6%**	59.4%	57.4%

Pedestrian Detection

- Dataset: **SED-PD:** as large as Caltech dataset
  - Randomly sample from TrecVid08-Dataset
  - Positive
    - 124,000 for training
    - 63,000 for validation
    - 210,000 for testing
  - Negative
    - Anything of non-positive
    - About 7 millions
  - Available: www.bupt-mcprl.net/datadownload.php

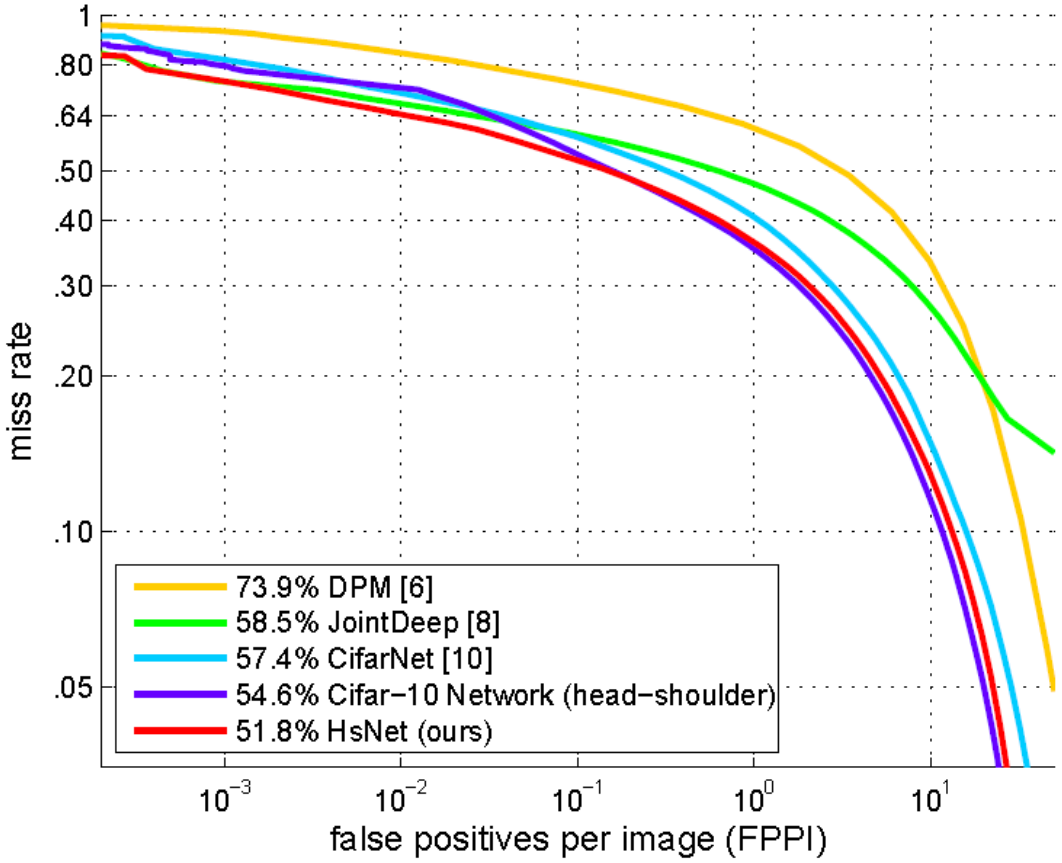
Pedestrian Detection

- **HsNet: A CNN Cascade Architecture**
  - **Input size: 32x32**



Pedestrian Detection

- **HsNet on SED-PD**



[10]J. Hosang, CVPR 2015
[6] P. Felzenszwalb, CVPR 2008
[8] P. Luo, CVPR 2014

Device	Model (second/frame)			
	Ours	CifarNet [10]	DPM [6]	JointDeep [8]
CPU	1.02	1.69	3.76	31.52
GPU	0.33	0.56	–	–

Outline

- Retrospective System Overview
- Pedestrian Detection
- Detected by CNN
  - Embrace and Pointing
  - ObjectPut and PersonRuns
- Pedestrian Tracking
- Detected by Trajectory Analysis
  - PeopleMeet and PeopleSplitUp
- Performance Evaluation
- Conclusion

Embrace, Pointing
ObjectPut, PersonRuns

- Regard four actions as the classification of static **key-poses**

Pointing

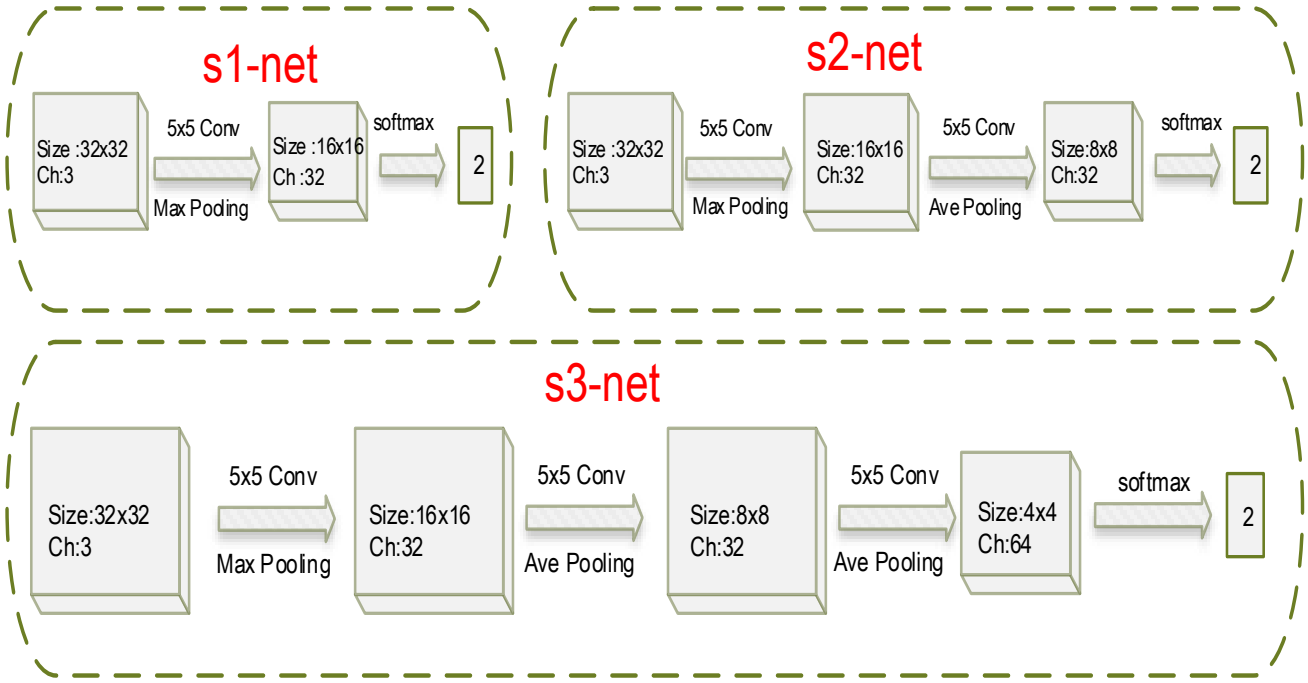
Embrace

ObjectPut

PersonRuns

Embrace and Pointing

- Method
  - **Apply the same CNN architecture of pedestrian detection**
  - Training samples are the pedestrian detection results with 1.5-fold expansion



Embrace and Pointing

Embrace:

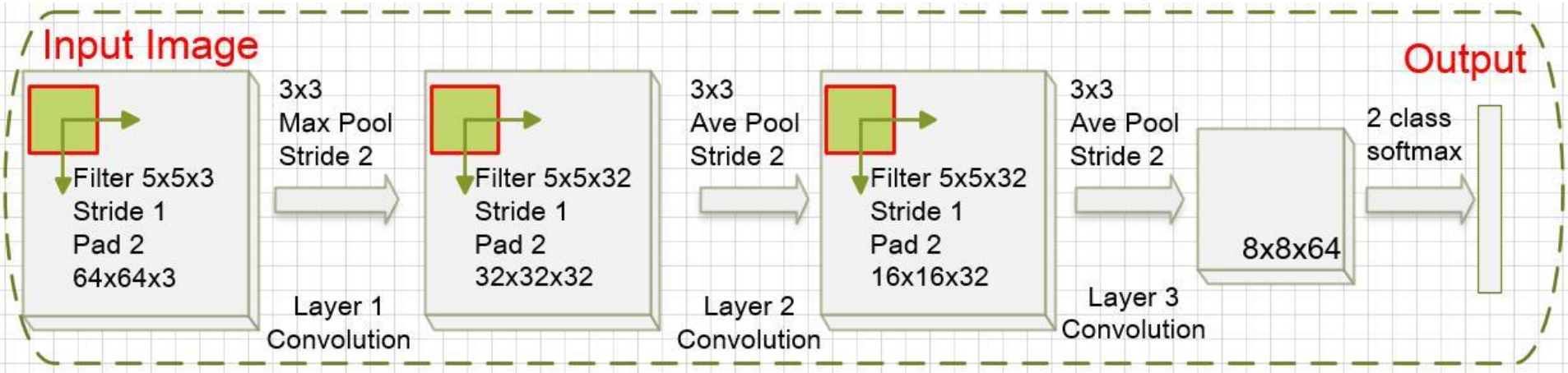
Year	ADCR	#CorDet	#FA	#Miss
2014	**0.8318**	**26**	**44**	**112**
2015	**0.7909**	**36**	**90**	**102**

Pointing:

Year	ADCR	#CorDet	#FA	#Miss
2014	**1.0027**	**21**	**57**	**774**
2015	**1.0040**	**16**	**42**	**778**

ObjectPut and PersonRuns

- Method
  - We train two CNNs to recognize them
  - The structure is similar to Cifar-10
  - The size of samples is 64x64



ObjectPut and PersonRuns

ObjectPut:

Year	ADCR	#CorDet	#FA	#Miss
2015	**1.0120**	**2**	**33**	**287**

PersonRuns:

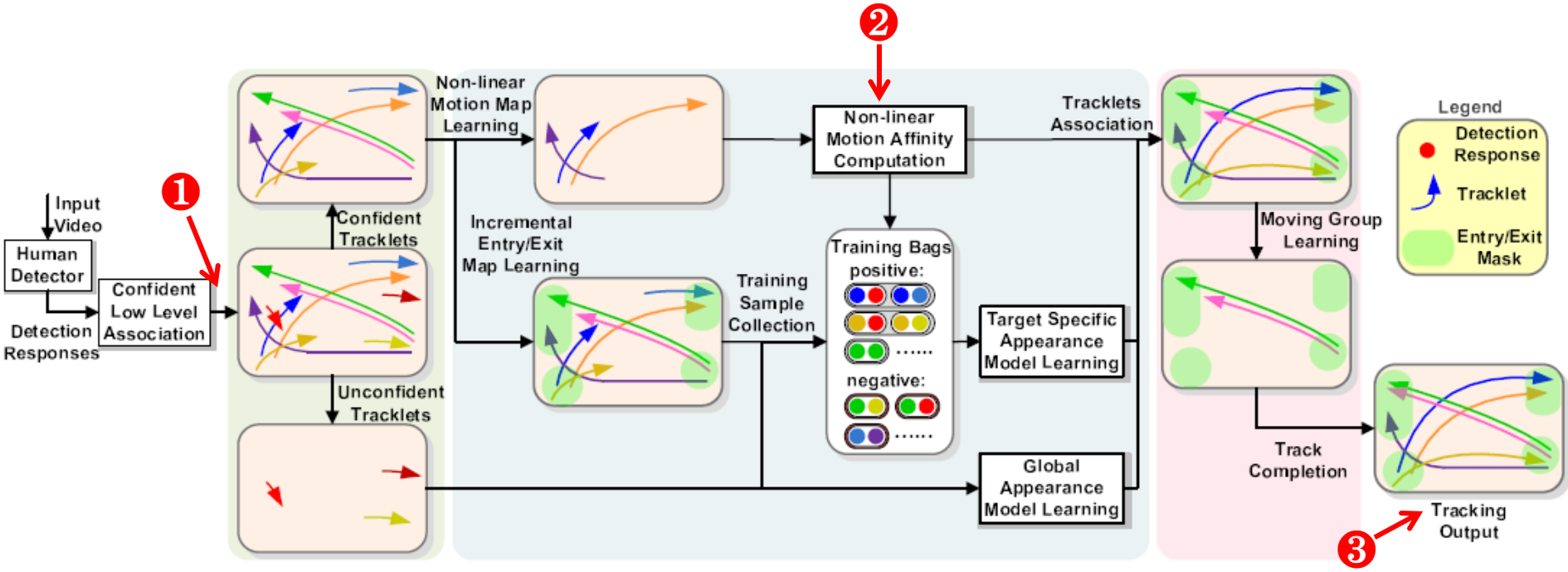
Year	ADCR	#CorDet	#FA	#Miss
2014	**0.9070**	**8**	**139**	**43**
2015	**0.9700**	**4**	**87**	**46**

Outline

- Retrospective System Overview
- Pedestrian Detection
- Pedestrian Tracking
- Detected by CNN
  - Embrace and Pointing
  - ObjectPut and PersonRuns
- Detected by Trajectory Analysis
  - PeopleMeet and PeopleSplitUp
- Conclusion

PeopleMeet & PeopleSplitUp： Trajectory Analysis based on Pedestrian Tracking

- Still followed this tracking method [Bo Yang. CVPR 2013]
  - Although we introduce Gaussian process regression instead of original quadratic function to improve tracking performance, the detection results are unsatisfactory.



Conclusion

- SED-PD: our dataset plays an important role.
- HsNet: a cascade-based CNN model improve the accuracy and speed of pedestrian detection which contributes to SED.
- In future work, smarter CNN-based method joint dense trajectory algorithm would be explored.

Thank you!