Online-sequential extreme learning machine based double thresholding strategy for human detection and tracking

KAIST, AIM Lab
Youngmin Park

Human detection/tracking on robot

- Human detection and tracking based on deep learning has been greatly improved
- Many robot services have become feasible based on the improvements
- Deep learning models Limit: lack of training data and difficulty of training
  The models require complex parameter tuning and large computational resources

Deep learning presumes 'a largely stable world'

Shallow model is also needed for robot

- Robots receive a **large amount of data** in **real time** from a **continuously changing environment** and have to handle **many requests** with **limited resources without human intervention**

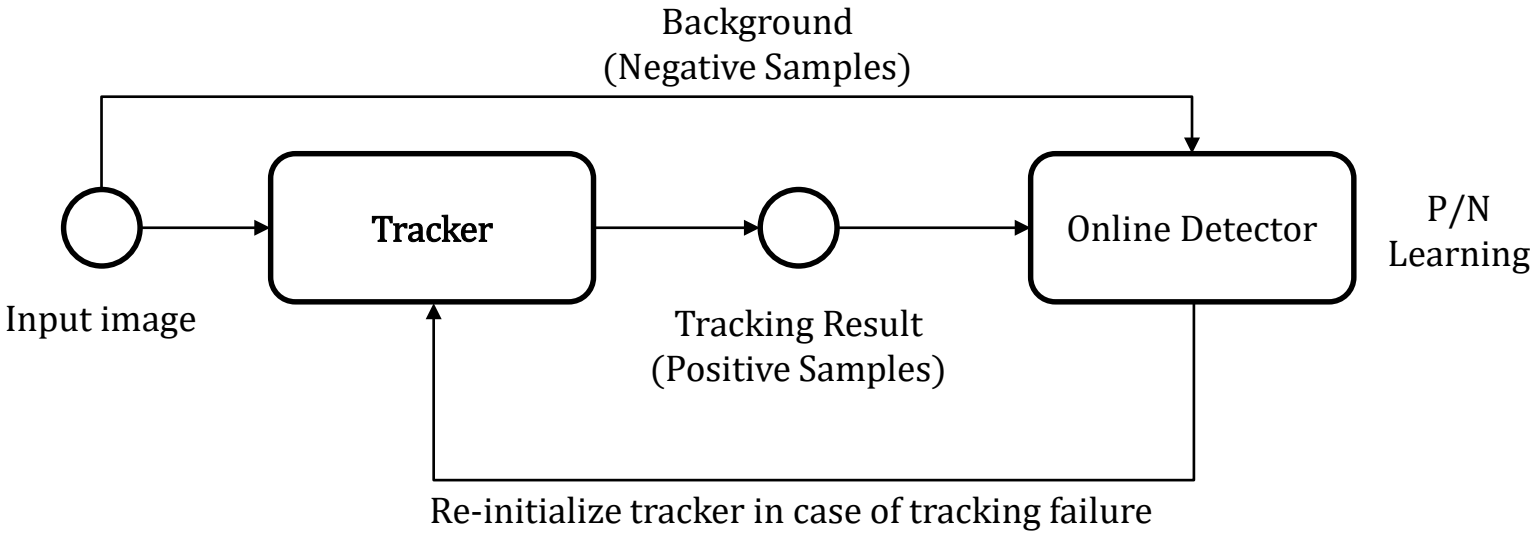
- Deep learning models are not suitable for the robot environment.

  A more adaptable and lightweight model is necessary for the robot service environment

- Deep Learning vs. Tracker vs. **Deep & Shallow learning**

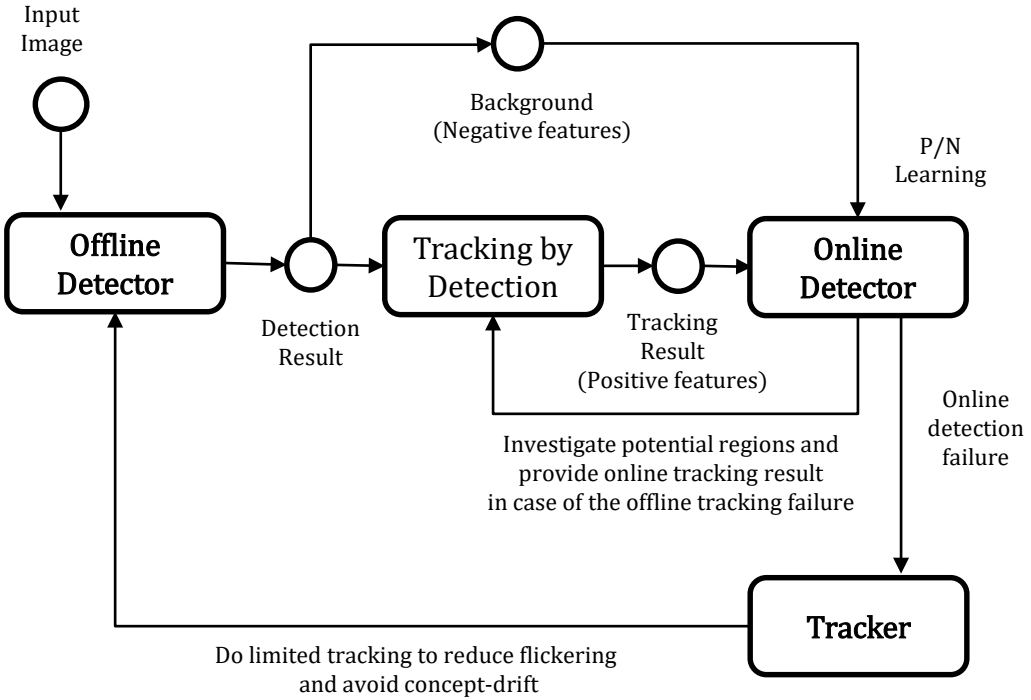
Tracking-Learning-Detection(TLD) is a representative baseline framework

- An online detector learns a target information from the environment based on tracking result



Proposal : Learning-Detection-Tracking(LDT)

- Shallow & Deep neural network is based.
- **High-recall and precision** from recent deep-learning based detection model.
- Tracker compensates detection results.



Proposal : Learning-Detection-Tracking(LDT)

- **Offline detector(YOLO)**

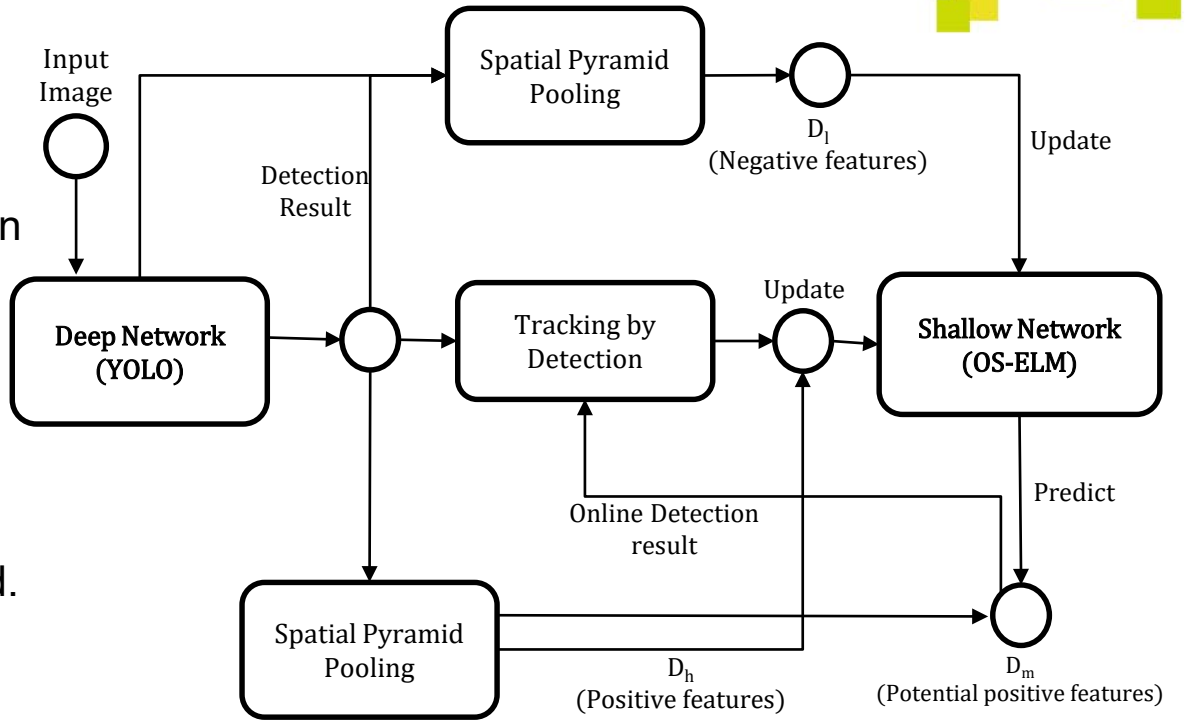
YOLO – You Only Look Once

Feature representation is shared to an online detector.

- **Online detector(OS-ELM)**

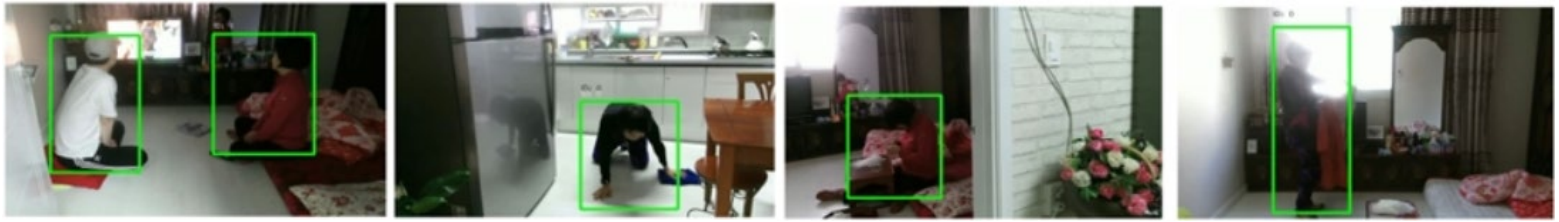
OS-ELM: Online-Sequential Extreme Learning Machine

Features of specific person is learned.



Experiment Result

- Real-world living lab environment (8,783 frames from 7 videos)
- Various situations : sitting, watching TV, hugging, cleaning, and wearing



Precision : +12%
Recall : -3%
MOTP : -

F1-score : +3.7%
MOTA : +14.2%

TABLE I
DETECTION AND TRACKING PERFORMANCE ON LIVING LAB DATASET

Method	Precision	Recall	F1-score	MOTP	MOTA
Baseline	0.685	0.855	0.76	0.782	0.462
Fine-tuned	0.763	0.866	0.81	0.786	0.597
Fine-tuned (TLD)	0.459	**0.888**	0.61	**0.787**	-0.16
Fine-tuned (DLT)	**0.854**	0.824	**0.84**	0.782	**0.682**

Conclusion

- Our proposed method outperformed the pure deep learning based approach and the sophisticated adaptive tracking approach such as TLD
- If additional information(ex. location information) is added to online learning, accuracy could be further improved
- Future works : using a variety of additional information for online detector