PRECISE REGRESSION FOR BOUNDING BOX CORRECTION FOR IMPROVED TRACKING BASED ON DEEP REINFORCEMENT LEARNING

Yifan Jiang, Hyunhak Shin, Hanseok Ko

School of Electrical Engineering, Korea University, Seoul, South Korea

Introduction

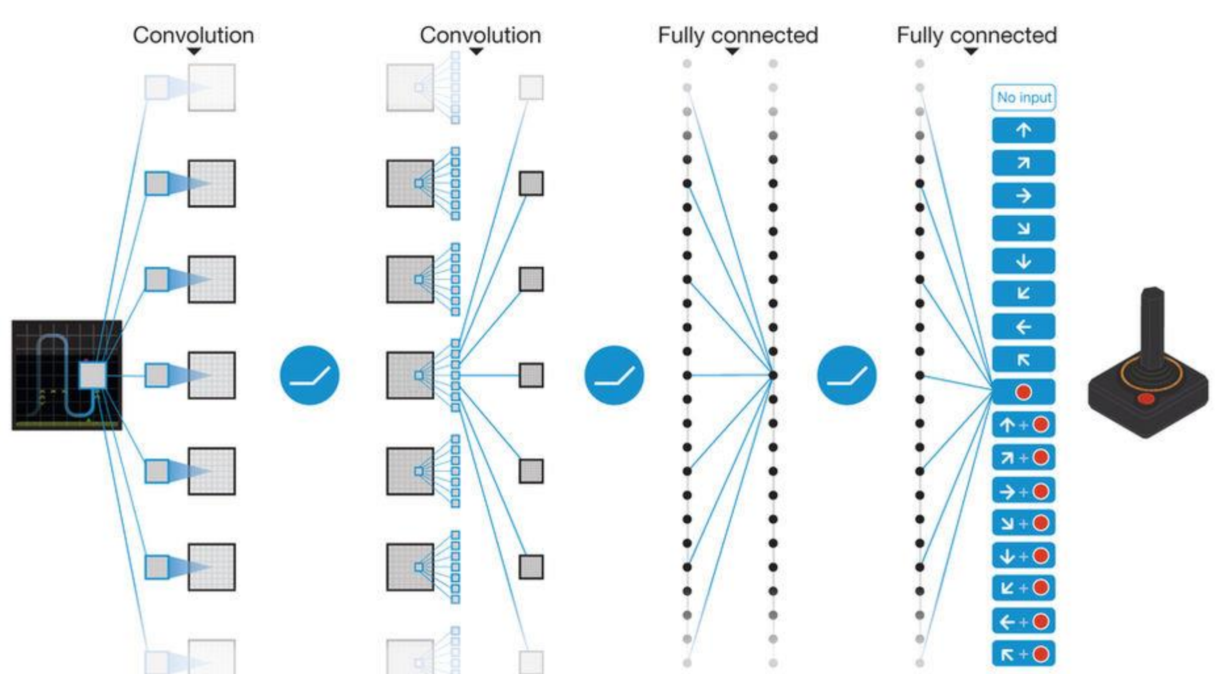
- **Motivation**
  - The state-of-art multi-object tracking accuracy is still limited to poor performance due to presence of complex scenes and frequent change of target appearance.
  - Poor performance of object tracking can be attributed largely by the object's encapsulation with inaccurate bounding box in the form of oversized, partial and false position as shown below

Oversized bounding box | Partial bounding box | False bounding box

- **Conventional Method**
  - There is almost no existed regression based correction method for multi-object tracking task;
  - Recently, regression approaches have been widely used in various object detection tasks [1, 2] for correcting the object detection results or finishing an object detection task.

Background

- **Deep Q-learning**



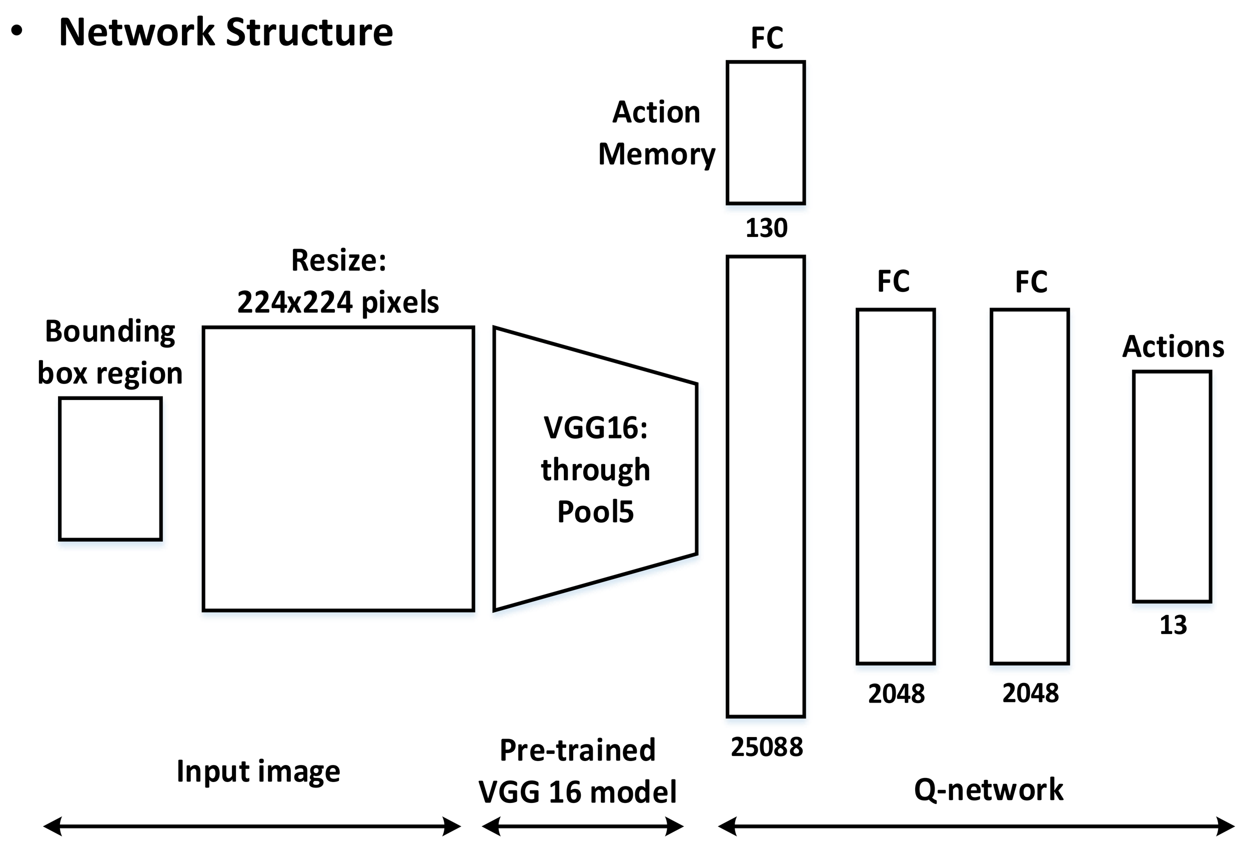
  - Deep reinforcement learning (DRL) is a series of enhanced algorithms which exploits deep learning theories to improve the original reinforcement learning methods. One typical algorithm is deep Q-learning (DQN) [3].
  - In DQN algorithm, a CNNs model is used to approximate the Q-table which indicates the future rewards that can be obtained by following a searching policy through all possible states. DQN learns about a state-action value function (Q-value) and represents Q-value by CNNs, then decides a sequence of actions following the output of CNNs model.

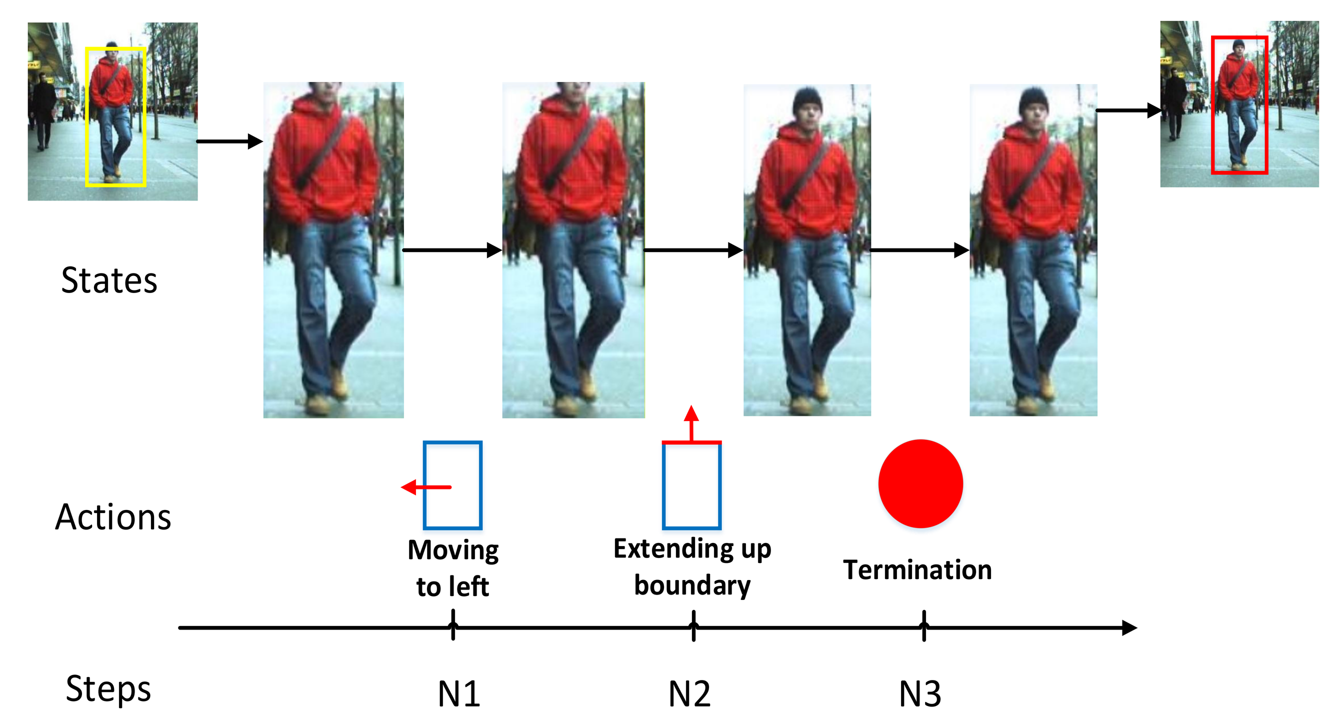
Proposed Method

- **Overview**
  - Error bounding box correction task essentially can be modeled as a framework of Markov decision process (MDP) because the resulting outcome is partly random and partly under the control of a decision maker.
  - We can exploit this hypothesis to model an agent to make the sequence of decisions.
  - We set a single bounding box region as environment (or observation), so that the agent can make actions to move the bounding box according to the environment.
  - Our proposed method follows a neighborhood search strategy, which starts from a random region near by previous target location and then adjusts position and size to correct target.

- **MDP Formulation**
  - Actions: There are 13 possible actions which can be categorized into movement actions (e.g. 4 actions), scale actions (e.g. 8 actions) and termination action (e.g. 1 action).
  - States: States in our work can be divided into two parts; feature vector and memory vector. The feature vector is the Pool5 layer feature map of VGG-16 from current bounding box region. The memory vector consists of the last 10 actions which the agent has already performed in search for an object.
  - Reward: Reward strategy of the proposed method closely follows the Caicedo and Juan's work [2]. To adjust the object tracking task, a specific case is needed. Hence, we set the threshold with a constant and $\tau = 0.9$, while other parameters stay the same as in [2].

- **Network Structure**



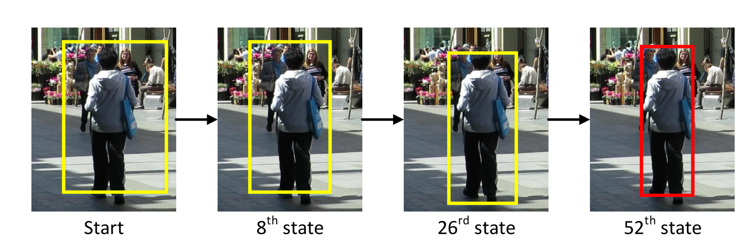


Experiment Results

- Performance comparison of original state-of-the-art multi-object tracking methods and methods with regression approach; while the best evaluation metric is in bold.

Object tracking method	Regression method	MOTA (%)	IDF1 (%)	MT (%)	ML (%)	HZ
AMIR15 [5]	OURS	**40.1**	46.0	**18.4**	**23.0**	0.7
	Girshick *et al* [4]	37.4	46.0	15.4	26.5	1.6
	He *et al* [1]	38.6	46.0	17.9	25.7	1.5
	None	37.6	46.0	15.8	26.8	**1.9**
HybridDAT [6]	OURS	**42.3**	47.7	13.6	**39.7**	3.1
	Girshick *et al* [4]	36.0	47.7	11.5	42.6	4.0
	He *et al* [1]	37.4	47.7	**13.8**	40.0	4.0
	None	35.0	47.7	11.4	42.2	**4.6**

- Example sequences of regression procedure.



ADL-Rundle-3 160th frame

PETS09-S2L2 314th frame

ETH-Linthescher 133th frame

Experiment Setting

- **Training procedure**
  - **Samples generation:** we follow a motion smoothness hypothesis which indicates the changes of object location and size obey Laplace distribution with mean of 0 and 1 respectively.
  - **Training strategy:** A $\epsilon - greedy$ policy is used to enlarge path searching range by randomly choosing actions. ϵ is initiated with 1 then decreased to 0.1 by steps of 0.05 every 5 epochs.
  - **DQN parameters:** Setting replay pool size as 1000 and discount factor $\gamma = 0.9$.
  - **Network parameters:** VGG 16 model is pre-trained by ImageNet database. Q-network is initiated randomly from a uniform distribution and trained with learning rate as 1e-6 and Adam optimizer. Finally, each target-specific model is trained with 100 epochs and batch size of 100.

- **Testing procedure**
  - Each regression iteration is limited in 100 steps;
  - Termination action appears within 100 steps is considered as a successful iteration. Otherwise, the iteration is considered as a failure.

- **Experiment**
  - **Dataset:** 2D MOT 2015;
  - **Evaluation metrics:**
    - ➢ MOTA (Precision), IDF1 (ID change), MT (Precision), ML (Precision) and HZ (Speed).

Conclusions

- In this paper, a precise bounding box regression approach to correct imprecise bounding box is proposed for improving tracking result of object tracking task.
- Our proposed method employed deep reinforcement learning algorithm to learn about how to explore for the optimal regression path between error bounding box and ground truth.
- Experimental results indicate that <u>the proposed regression method can correct error bounding box effectively and definitely increase the tracking accuracy of state-of-the-art object trackers.</u>

References

[1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in European Conference on Computer Vision. Springer, 2014, pp. 346–361.

[2] Juan C Caicedo and Svetlana Lazebnik, "Active object localization with deep reinforcement learning," in Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 2488–2496.

[3] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al.: Human-level control through deep reinforcement learning. Nature 518(7540) (2015) 529-533.

[4] Ross Girshick, "Fast r-cnn," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440–1448..

[5] Amir Sadeghian, Alexandre Alahi, and Silvio Savarese, "Tracking the untrackable: Learning to track multiple cues with long-term dependencies," arXiv preprint arXiv:1701.01909, 2017..

[6] Min Yang, Yuwei Wu, and Yunde Jia, "A hybrid data association framework for robust online multi-object tracking," arXiv preprint arXiv:1703.10764, 2017.

yfjiang@ispl.korea.ac.kr *given_jiang@hotmail.com*