**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 VGG16 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 \( \eta = 0.9 \), while other parameters stay the same as in [2].

**Network Structure**

**Proposed Method**

**Conclusion**
- 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 the proposed regression method can correct error bounding box effectively and definitely increase the tracking accuracy of state-of-the-art object trackers.

**References**