Real time object detection framework with deep Reinforcement learning

Mrs Sonal Tiwari *, Dr. Shailja Sharma **, Dr. Saurabh Tiwari ***

* Research Scholar, Department of Computer Science and Engineering RNTU Bhopal
** Associate Professor, Department of Computer Science and Engineering RNTU Bhopal
***Sr. Lecturer, Government S. V. Polytechnic College Bhopal

Abstract- Object Detection with optimized tracker is a complex process because it faces many problems such as occlusion, blur image, background clutter, fast motion and many obstacles. Tracker which is based on Reinforcement learning can solve this problem, RODRLM (Real Time Object Detection Deep Reinforcement Learning Model) based on reinforcement learning can track the object with great improvement of previous tracker. Although ADNet has some limitation in optimal action selection, and suffer from inefficient tracking .So some improvement is supposed to improve ADNet to enhance the tracking accuracy and efficiency. Firstly, the multi-domain training is incorporated into ADNet to further improve the feature extraction ability of its convolution layers. Then, in the reinforcement learning based training phase, both the selection criteria for optimal action and the reward function are redesigned separately to explore more appropriate action and eliminate useless action. an effective online adaptive update strategy is proposed to adapt to the appearance changes or deformation of the object during actual tracking.

Index Terms- Computer Vision, Deep Learning, Reinforcement Learning, Visual Tracking.

I. Introduction

Present paper gives a concept to detect object real time with effective quality and can implement in any device with limited resources. As it is known to us, letting the agent possess the ability to continuously learn and adapt from limited experience in non-stationary environments is an important milestone on the path towards general intelligence. supervised learning with outstanding outcome in image classification [1]-[5], image segmentation [6]-[9] and object detection [10]-[13], reinforcement learning with excellent strategy-making ability in an unstable environment [14]-[17], and meta-learning that can quickly adapt to a new task with a small amount of samples [18]-[22]. with the development of machine learning, numerous learning networks have been widely used in visual tracking field [23]-[26]. In 2017, for the first time,Yun et al.[24] proposed a novel tracker controlled by the designed action-decision network (ADNet), which is radically different from the existing trackers. In this algorithm, the tracker is defined as an agent of which goal is to capture the object with a bounding box. First, ADNet is trained separately by using supervised learning and reinforcement learning, and then the agent is able to decide sequential action until finding the location of the object in each video frame. Compared with existing trackers based on deep networks, ADNet has achieved better results in tracking accuracy and efficiency.

However, there are some following deficiencies in ADNet.There are insufficient tracking training sequences to fine-tune the network for specific tasks through pre-trained or transfer learning, like large-scale classied networks such as ImageNet [27], ResNet [1], and VGGNet [2].

The primary contributions of this paper are summarized as follows. (1) The reinforcement learning based training method is improved by redesigning the decision criteria for optimal action to explore more appropriate action and redesigning the reward function to eliminate useless action. (2) A meta-learning based online adaptive update scheme for model parameters is proposed to adapt to the appearance changes or deformation of the object in the actual tracking. (3) The multi-domain training is incorporated into the ADNet model to further enhance the ability to learn the generic representation of different objects.

II. RELATED WORK

The recent object detection algorithms that apply deep learning methods to overcome the problem, each method has used deep learning in a different way to enhance effectiveness and efficiency. As the methods evolve, they move closer to the idea of end to- end learning.

A. VISUAL OBJECT TRACKING

Visual object tracking can be generally divided into two categories, including a generative-based approach and a discriminant-based approach .Among them, the generative based approach usually extracts the features and learns them to generate a model representing the appearance of the object, and then the region that best matches the model is considered as the object in the image. While the discriminant- based approach, from a mathematical point of view, poses the tracking problem as a binary classification problem, and its task is to distinguish the object from the surrounding background.

Convolutional neural network (CNN) has been proven with outstanding performance in a wide range of computer vision applications [2]-[7]. Despite this, the early CNN application [33] in tracking suffered from the data deficiency problem for training its network. To solve the data insufficiency, the transfer methods [9], [34] were proposed by using the pre-trained classification dataset (such as Image Net [27]). However, these methods still have limitations due to the gap between image classification and object tracking. The recently proposed methods [23], [35] tried to defeat this gap by training the networks with a large number of tracking training datasets [36]. Tao et al. [35] designed a Siamese deep neural network to learn a matching function, which is used to pursue the most similar candidates in a new frame through using only the original observation of the object from the rst frame. Cui et al. [37] proposed a Recurrently Target-attending Tracker (RTT) by using multi-directional recurrent neural network to identify and exploit reliable patches that facilitate the entire tracking process.

B. REINFORCEMENT LEARNING

Reinforcement learning, as a common method of training agents, is able to accomplish many complex tasks. For reinforcement learning, the strategy-making ability of the agent is trained by maximizing the reward function. In recent years, the development trend of reinforcement learning is to combine the perception ability of deep learning with the strategy-making ability of reinforcement learning. Since the combined learning algorithm of reinforcement learning and deep learning has been proposed, it has successfully solved many complex problems in challenging tasks that need to perceive high-dimensional raw input data and strategy determine, such as [16], [24].

C. META-LEARNING

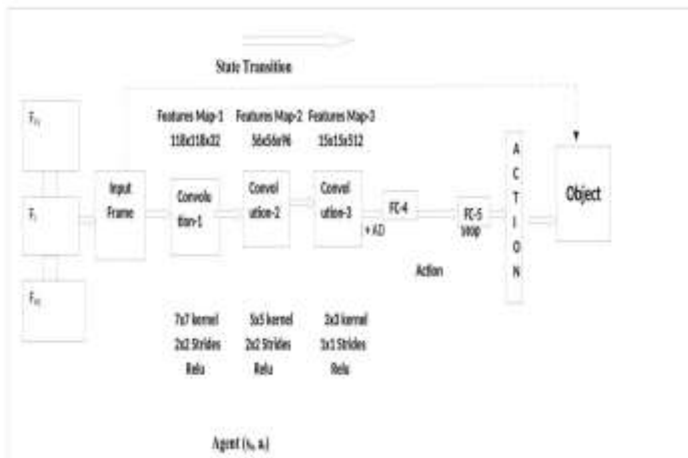
Meta-learning is the product of the successful development of deep learning, which enables the model to learn from previous experience and quickly adapt to new tasks [18], [19], [25]. A variety of meta-learning based methods are available, such as memory-based meta-learning method [20], which establishes a connection between the label and the input so that subsequent inputs can obtain relevant data through external memory and perform comparison to achieve better prediction. Adam et al. [20] developed a model-agnostic meta-learning (MAML) algorithm, in which meta-learning is used to train the parameters of the model on various learning tasks with only a small amount of training data from the current task, and then MAML is able to effectively solve new learning tasks. Finn et al. [21] re-derived MAML for multi-task reinforcement learning from a probabilistic perspective and extended it to dynamically changing tasks, which enables the trained agent to continue to learn in nonstationary environments and optimize its own strategy. Park and Berg [25] improved state-of-the-art on-line trackers based on deep network and utilized the meta-learning-based method to adjust deep network for tracking using offline training, so that the network can quickly be adapted to a specific object in future frames.

III. PROPOSED TRACKER

Proposed tracker is developed to increase tracking accuracy and efficiency. The multi-domain training is soul of algorithm which makes the combination of training effective for proposed work. Selection criteria for optimal action and the reward function are redesigned separately to explore more appropriate action and eliminate useless action. an effective online adaptive update strategy is proposed to adapt to the appearance changes or deformation of the object during actual tracking. Experimental results demonstrate that the proposed tracker has advantages over ADNet and other techniques in terms of accuracy and efficiency.

A. RODRLM (real time object detection deep reinforcement learning model)

The proposed model avoids the robustness and difficulties of ADNet model by improving the convolution layer features map. firstly video is divided into multiple frames ,these frames are the input frame ,each frame is the combination of multiple patches .now each frames goes to the convolution layers ,convolution layer have been used for extract the feature of frame ,enhance edge .this layer Convolution layers receive as input an image I(m-1) (with Lm channels) and compute as output a new image I(m) . The output of each step is known as feature Map .Depending upon the feature extraction tracker take the action .Overall Model is divided into three parts.

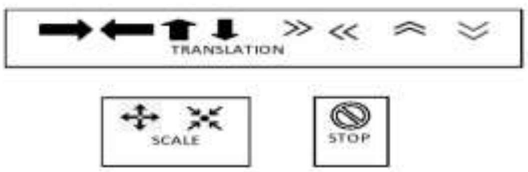


RODRLM

To improve the robustness and real-time performance of the tracker, research has been improved from three aspects. Firstly, the use of multi-domain training instead of supervised learning based training enables the tracker to learn the shared representation of different objects in the various training sequences, which allows the model to possess the ability to make single-step action decision. Secondly, the policy gradient based reinforcement learning is improved so that the tracker can capture the object by selecting more appropriate action and eliminating the useless action. Thirdly, the meta-learning based online adaptive update scheme is proposed to pursue the optimal parameters for the network so that the tracker can quickly adapt to new tracking tasks.

B. Network structure

The architecture of the multi-domain training network, which is used to predict a single-step action for object tracking, is shown in Fig. 2. It receives 112 x 112 three-channel patches as the input, and has six hidden layers, including three convolutional layers (conv1-3) and two fully connected layers (fc4-5). Among them, the last fully connected layer (fc5) predicts the probability corresponding to 11 kinds of action (As shown in Fig. 3, these action can be divided into three categories: translation, scale change, and stop, where, the translation moves consist of four directional moves, {left, right, up, down}, and also include their two times larger moves; the scale changes are defined {scale up, scale down} [24]). Moreover, all the convolutional layers (conv1-3) and the two fully connected layers (fc4-5) form the shared layers, and the fully connected layer (fc5) consists of k branches of domain-specific layers.



Action space

IV. TRAINING

The training samples are first generated according to the image and the ground truth of each frame in training datasets, which is similar to the supervised learning based training process of the ADNet [24]. Since the training datasets provide video frames f and the ground truth gt , Gaussian noise is randomly added to the ground truth to produce a series of ``boxes" b i that deviate from the ground truth, and then the image patches pi cropped with the ``box" bi are used as training patches in a frame image f,

$$p_{i=} \; c(b_{i,}f) \; , \qquad (1)$$

where, C (.) represents the function which crops the patch p i from the frame f with the ``box" b i and resizes pi to match the input size of the network; the ground truth is gt denote the optimal center location, and $\hat{w}$ and $\hat{h}$ denote the width and height of the optimal tracking bounding box,

$$\hat{y}_i = \arg\max_{\alpha} \mathrm{IOU}(f(b_i, a), gt), \qquad (2)$$

where , f $(b_{i,}, a)$ denotes the moved patch from "box" b_i by the action a and $\mathrm{IOU}(f(b_i, a), gt)$ represents overlap rate of the "box" location and the ground truth gt of the object with intersection-over -union criterion after the action a moving.

This completes the entire process required for widespread of research work on open front. Generally all International Journals are governed by an Intellectual body and they select the most suitable paper for publishing after a thorough analysis of submitted paper. Selected paper get published (online and printed) in their periodicals and get indexed by number of sources.

The problem settings for the reinforcement learning in the visual tracking are as follows: State. The state s is the patch obtained by cropping the image *f* with the current tracking bounding box, which is defined as (1). Action$\{ a_{1,\dots\dots,} a_{11} \}$, as shown in Fig. 3. Reward. The reward function we designed is different from the one in ADNet. In the original ADNet, the frames $\{\{f_1, \dots, f_n\}$ contained in a certain piece of the training sequence are used to calculate the reward according to the tracking result of the first frame f_1 and the last frame f_n, and the calculated reward is treated as the reward for each action of each frame $\{f_{1,} \dots \dots f_n\}$,Although this is a relatively easy way, there are some problems.

To deal with the issues raised here , we allocate a reward to each frame of the training sequence, and the designed reward function is assigned by,

$$r = \begin{cases} \lambda[\mathrm{IOU}(f(b', a'), gt) - \mathrm{IOU}(f(b, a), gt)], \\ \qquad \text{if } a' \neq \text{stop} \\ +1 \;, \quad \text{if } a' = \text{stop and IOU } (f(b, a), gt) > 0.7 \\ -1 \;, \quad \text{if } a' = \text{stop and IOU } (f(b, a), gt) < 0.7 \end{cases} \qquad (3)$$

where,a' and b' represent the action and the bounding box of the next moment, respectively. λ ε 2 [1; 10] denotes a constant.

Considering that there are many action dimensions and a large searching space in the problem settings of the reinforcement learning based training, the policy gradient algorithm [28] is employed to train the model. For a training sequence, 200 consecutive frames are randomly selected, and then the ground truth gt of the first frame is used as the initial ``box" bi and the image is cropped with the ``box" to obtain the patches as the initial state S . Finally, take S as the input of the network to calculate the conditional probability of 11 action.

The training process of exploring more appropriate action is as follows

Step 1, the ``box" is shifted or scaled to obtain a new ``box" and a new ``patch" according to the selected action, as illustrated in Fig. 4

Step 2, the new ``box" and the new ``patch" are used as the input to the network to decide the next appropriate action until the stop action is selected or the number of action reaches a certain threshold.

Step 3, end the current frame and calculate its reward, and then proceed to the next frame.

Repeat the above steps until all frames are over. Then, calculate the gradient and update the network parameters,

$$w_{rl\leftarrow} w_{rl} + \sum_l^L \sum_l^T \nabla_{rl} \log p(a_{t,} | s_t, l; W_{rl}) * r_t, l, \qquad (6)$$

V. Result and Analysis

The performance of the proposed approach is evaluated based on the challenging online object tracking benchmark (OTB) [41] in the experiment.VOT2013 [42],VOT2014 [43], VOT2015 [44] datasets were selected as training samples, and VOT-100 dataset [41] with 100 video sequences (includingOTB-50) was selected as the test samples.

Comparison of evaluation indicators of different trackers

To further validate the effectiveness of our tracker, 9 stateof- the-art trackers, including ADNet [24], CCOT [32], MDNet [23], MCPF [34], ECO-HC [45], DCFNet [26], MEEM [31], DSST [46] and KCF [30], are selected to compare with our tracker. Table 1 show the precision and success performance based on center location error and overlap ratio, respectively.

Precision describes as accuracy of measurement and intersection of union tells how much object actually detected in particular action. Calculation of both gives the accuracy of object detection .following table compares proposed technique with other existing research and gives the results and with these results we can say that proposed research leads object detection in smarter way.

	PROPOSED	ADNET	CCOT	MDNET	ADNET	MCPF	ECO-HC	DCFNET	MEEM	KCF	DSST
Precision	87.1%	86.8%	86.9%	86.3%	84.4%	84.5%	82.7%	77.8%	75.4%	66.8%	66.5%
IoU	.655	.651	.653	.648	.628	.613	.625	.598	.515	.462	.502

Precision and IoU performance

As table shows proposed tracking model works well comparative to others. Another important factor fps that is frame per seconds plays better role for defining the efficiency of model .following table shows comparison with other technique

	PROPOSED	ADNET	CCOT	MDNET	ADNET	MCPF	ECO-HC	DCFNET	MEEM	KCF	DSST
FPS	4.1	3.9	<1	<1	3.0	<1	16.0	29.0	20.1	124.6	16.1
GPU	0	0	0	0	0	0	0	0	0	0	0

Comparison based on FPS and GPU

above two results shows that proposed technique is more promising than other.

VI. CONCLUSION

In this paper ,limitation of existing models is overcome through the multi-domain training ,improve the criteria of the most favorable action selection and three convolution network layer ,with increase patches and reinforcement learning based training and utilizing meta -learning for the parameters of the network during object tracking to improve the efficiency .For video 128 training patches and 8 frames are randomly selected for each patch size of frame .Proposed tracker consist fewer layers in network and its computational complexity is lower as compared to existing deep networks for tracking. According to the experimental results, the speed of the proposed tracker improves 4.2 frames per second, which is 38.5 % faster than the original ADNet. Proposed tracker is compared with many existing network tracker, and we found much better result as previous tracker in precision and success rate.

REFERENCES

[1] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, ``Squeeze-and-excitation networks," IEEE Trans. Pattern Anal. Mach. Intell. , to be published. doi: 10.1109/TPAMI.2019.2913372.

[2] K. He, X. Zhang, S. Ren, and J. Sun, ``Deep residual learning for image recognition," in Proc. CVPR , Las Vegas, NV, USA, Jun. 2016, pp. 770-778.

[3] K. Simonyan and A. Zisserman, ``Very deep convolutional networks for large-scale image recognition," in Proc. ICLR, 2015, pp. 114.

[4] S. Ioffe and C. Szegedy, ``Batch normalization: Accelerating deep network training by reducing internal covariate shift," in Proc. ICML, Lille, France, 2015, pp. 448456.

[5] O. Ronneberger, P. Fischer, and T. Brox, ``U-net: Convolutional net- works for biomedical image segmentation," in Proc. MICCAI . Cham, Switzerland: Springer, 2015, pp. 234-241.

[6] K. He, G. Gkioxari, P. Dollar, and R. Girshick, ``Mask R-CNN," IEEE Trans. Pattern Anal. Mach. Intell., to be published. doi: 10.1109 TPAMI.2018.2844175.

[7] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, ``DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834-848, Apr. 2018.

[8] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, ``DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834848, Apr. 2018.

[9] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, ``DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834848, Apr. 2018.

[10] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ``You only look once:Unied, real-time object detection," in Proc. CVPR, Las Vegas, NV, USA,Jun. 2016, pp. 779-788.

[11] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, ``Feature pyramid networks for object detection," in Proc. CVPR, Honolulu, HI, USA, Jul. 2017, pp. 936-944.

[12] L. Pinto, J. Davidson, R. Sukthankar, and A. Gupta, ``Robust adversarial reinforcement learning," in Proc. ICML, Sydney, NSW, Australia, 2017, pp. 2817-2826.

[13] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, and D. Hassabis, ``Mastering the game of go without human knowledge," Nature, vol. 550, no. 7676, pp. 354359, Oct. 2017.

[14] K. Clark and C. D. Manning, ``Deep reinforcement learning for mentionranking coreference models," in Proc. EMNLP, Austin, TX, USA, 2016, pp. 1-7.

[15] O. Vinyals, C. Blundell, T. Lillicrap, K. kavukcuoglu, and D. Wierstra, ``Matching networks for one shot learning," in Proc. NIPS, Barcelona, Spain, 2016, pp. 3630-3638.

[16]] O. Vinyals, C. Blundell, T. Lillicrap, K. kavukcuoglu, and D. Wierstra, ``Matching networks for one shot learning," in Proc. NIPS, Barcelona, Spain, 2016, pp. 36303638.

[17] J. Snell, K. Swersky, and R. Zemel, ``Prototypical networks for few-shot learning," in Proc. ICLR, Toulon, France, 2017, pp. 40774087.

[18] M. Al-Shedivat, T. Bansal, Y. Burda, I. Sutskever, I. Mordatch, and P. Abbeel, ``Continuous adaptation via meta-learning in non stationary and competitive environments," in Proc. ICLR, Vancouver, BC, Canada, 2018, pp. 1-21.

[19] H. Nam and B. Han, ``Learning multi-domain convolutional neural networks for visual tracking," in Proc. CVPR, LasVegas,NV, USA, Jun. 2016, pp. 4293-4302.

[20] H. Nam and B. Han, ``Learning multi-domain convolutional neural networks for visual tracking," in Proc. CVPR, LasVegas,NV, USA, Jun. 2016, pp. 42934302.

[21] S. Yun, J. Choi, Y. Yoo, K. Yun, and J. Y. Choi, ``Action-decision networks for visual tracking with deep reinforcement learning," in Proc. CVPR, Honolulu, HI, USA, Jul. 2017, pp. 1349-1358.

[22] S. Yun, J. Choi, Y. Yoo, K. Yun, and J. Y. Choi, ``Action-decision networks for visual tracking with deep reinforcement learning," in Proc. CVPR, Honolulu, HI, USA, Jul. 2017, pp. 13491358.

[23] E. Park and A. C. Berg, ``Meta-tracker: Fast and robust online adaptation for visual object trackers," in Proc. ECCV, Munich, Germany, 2018, pp. 587-604.

[24] E. Park and A. C. Berg, ``Meta-tracker: Fast and robust online adaptation for visual object trackers," in Proc. ECCV, Munich, Germany, 2018, pp. 587604.

[25] Q. Wang, J. Gao, J. Xing, M. Zhang, and W. Hu, ``DCFNet: Discriminant correlation lters network for visual tracking," 2017, arXiv:1704.04057. [Online]. Available: https://arxiv.org/abs/1704.04057

[26] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, ``ImageNet large scale visual recognition challenge," Int. J. Comput. Vis.,vol. 115, no. 3, pp. 211-252, Dec. 2015.

[27] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, ``ImageNet large scale visual recognition challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211252, Dec. 2015.

[28] R. J. Williams, ``Simple statistical gradient-following algorithms for connectionist reinforcement learning," Mach. Learn., vol. 8, nos. 34, pp. 229-256, May 1992.

[29] H. Li, Y. Li, and F. Porikli, ``Robust online visual tracking with a single convolutional neural network," in Proc. ACCV. Cham, Switzerland: Springer, 2014, pp. 194-209.

[30] R. Tao, E. Gavves, and A.W. M. Smeulders, ``Siamese instance search for tracking," in Proc. CVPR, Las Vegas, NV, USA, Jun. 2016, pp. 1420-1429.

[31] Y.Wu, J. Lim, and M. H.Yang, ``Object tracking benchmark," IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9, pp. 1834-1848, Sep. 2015.

[32] Z. Cui, S. Xiao, J. Feng, and S. Yan, ``Recurrently target-attending tracking," in Proc. CVPR, Las Vegas, NV, USA, Jun. 2016, pp. 1449-1458.

[33] H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1985, ch. 4.

[34] B. Smith, "An approach to graphs of linear forms (Unpublished work style)," unpublished.

[35] E. H. Miller, "A note on reflector arrays (Periodical style—Accepted for publication)," *IEEE Trans. Antennas Propagat.*, to be published.

[36] J. Wang, "Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication)," *IEEE J. Quantum Electron.*, submitted for publication.