International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 5 May 2023, pp: 244-246 www.ijaem.net ISSN: 2395-5252

Deep Learning Approaches for Target Tracking

Arsh Dayal
Manipal Institute of Technology
Computer Vision [ICT - 4031]

Date of Submission: 01-05-2023 Date of Acceptance: 10-05-2023

ABSTRACT

The deep learning approach to target tracking combines object detection and object tracking using deep reinforcement learning. It deep involves a deep neural network that represents the policy and a reinforcement learning algorithm that trains the policy to maximize the reward signal. The deep neural network maps the current state of the target and environment to track actions while the reinforcement learning algorithm adjusts the network parameters.

Deep Learning Approaches for Target Tracking

Deep learning approaches for visual object tracking have become increasingly common due to the success of deep neural networks in various computer vision tasks. Object tracking is noticing and trailing a moving object in a video sequence [1]. It is crucial in multiple applications, such as surveillance, human-computer interaction, and autonomous systems [1].

The traditional approach to object tracking involves using hand-crafted features, such as optical flow and scale-invariant feature transform (SIFT), to detect and track the object. However, these approaches are limited by their dependence on hand-crafted features, which are only sometimes robust to changes in object appearance and illumination.

On the other hand, the basis of deep learning approaches for object tracking is end-to-end learning. Intuitive understanding of these features and the search are from the data. It allows for more robust and flexible monitoring, especially in challenging scenarios where the object's appearance and motion change over time.

This paper will discuss the deep learning approaches for visual object tracking, including object detection and tracking, deep reinforcement learning, Siamese networks, and object tracking using correlation filters.

A. "Object Detection And Tracking"

The "object detection and tracking" approach involves two steps: object detection and object tracking [4]. Experts use a convolutional neural network (CNN) in the object detection step to detect and localize objects of interest in each video sequence frame. The CNN takes the raw image data as input and outputs the location and size of the things in the image.

Once the objects have been detected and localized, the object tracking step begins. This step involves associating the objects detected in the current frame with those caught in previous frames. Researchers typically use techniques such as Kalman filters, particle filters, and recurrent neural networks (RNNs). RNNs are particularly well suited for this task because they can model the temporal dynamics of the object motion [4].

One of the advantages of the object detection and tracking approach is that it can handle multiple objects in a scene [4]. Professionals use the object detection step to detect objects in the initial frame, eliminating the need for manual initialization.

B. Deep Reinforcement Learning

The deep reinforcement learning approach leverages the power of deep neural networks to learn an optimal policy for target tracking in a given environment [4]. In this approach, experts train a deep neural network to make real-time tracking decisions by interacting with the environment and learning from its mistakes [4]. Professionals can view the deep reinforcement learning approach to target tracking as combining two key components: a deep neural network representing the policy and a reinforcement learning algorithm that trains the policy [4].

The deep neural network is responsible for mapping the current state of the target and the environment to a set of tracking actions. At the same time, the reinforcement learning algorithm adjusts the parameters of the network to maximize the reward signal that measures the tracking system's performance.

One of the key advantages of the deep reinforcement learning approach to target tracking is that it allows for real-time adaptation to changes in the target motion and the environment; This is particularly important in scenarios where the target motion is unpredictable and the atmosphere is dynamic [4].

C. Siamese Networks

The siamese network consists of two branches, each of which takes a pair of consecutive frames as input and outputs the similarity score between the two frames [4]. Professionals use the similarity score to determine the success of tracking the object from the first frame to the next. Experts train the network using a contrastive loss function that encourages the network to output a high similarity score for pairs containing the same object and a low similarity score for pairs of frames containing different things [4].

Disadvantages of the Approaches

A. "Object Detection and Object Tracking"

One of the main disadvantages of this approach is that it requires large amounts of annotated training data to train the CNNs and RNNs [4]. The quality of the tracking performance is highly reliant on the standard and quantity of the training information, and it can be difficult and time-consuming to collect and annotate large amounts of data. This approach is computationally expensive, especially with high-resolution videos and real-time requirements.

B. Deep Reinforcement Learning for Target Tracking

One of the main disadvantages of this approach is the difficulty in defining the reward signal that measures the tracking system's performance [4]. The reward signal should accurately reflect the desired behavior of the tracker. Still, it can be challenging to design a reward function that captures all relevant factors, such as accuracy, robustness, and computational efficiency. Additionally, this approach can be sensitive to the choice of hyperparameters, such as the learning rate, discount factor, and exploration strategy. It may require fine-tuning and tuning to work well in practice [4].

It's also important to note that the deep learning approaches for target tracking are still an active area of research, and there is still much work to be done to overcome these limitations and improve the performance of these methods in real-world scenarios. Despite these limitations, deep learning approaches represent a promising direction for developing new and more effective strategies for target tracking.

C. Siamese Network

Siamese networks can be computationally expensive and require much memory to store the parameters of the network [4]. It can be a challenge for real-time tracking, especially when dealing with high-resolution videos [4]. Also, Training a Siamese network for target tracking can be difficult, as it requires large amounts of annotated training data and can be sensitive to the choice of network architecture and hyperparameters [4]. Additionally, experts must train the network to handle various target appearances and motion patterns, which can be challenging.

The Best Approach

The best approach for target tracking using deep learning will depend on the specific requirements and constraints of the application. Each deep learning approach has its advantages and weaknesses. The method option is contingent on aspects such as the complexity of the target motion, the type of environment in which the tracking is taking place, and the computational resources available.

For instance, CNNs are well-suited for object detection tasks and can effectively detect and localize targets in each video sequence frame. However, they may not be as effective for tracking functions as they do not archetypal the "temporal dynamics" of the target motion. RNNs can be used to prototype the "temporal dynamics" of the target motion and are well-suited for target tracking tasks. However, they can be computationally expensive and require large amounts of annotated training data. Siamese networks are a popular approach for target tracking, as experts train them to learn the similarity between target appearances in different frames. However, they may not be as robust as other deep learning approaches in handling target appearance and motion changes.

CONCLUSION

Overall, deep learning has shown promise in the field of target tracking. However, many challenges still need to be addressed, such as dealing with occlusions, handling target disappearances, and cluttered scenes. Additionally, deep learning-based approaches to target tracking can be computationally expensive, which limits their practicality in a resource-constrained environment. Also, the best strategy for target tracking using deep learning will depend on the specific requirements and constraints of the application.

REFERENCES

[1] Kalake, L., Wan, W., & Hou, L. (2021). Analysis based on recent deep learning approaches applied in real-time multi-object tracking: a review. IEEE Access, 9, 32650-32671. https://ieeexplore.ieee.org/abstract/document/9359743

[2] N. An & W. Qi Yan, (2021). Multitarget tracking using Siamese neural networks. ACM Transactions on Multimedia Computing Communications and Applications, 17(2s), 1-16. https://dl.acm.org/doi/abs/10.1145/3441656

[3] P. Dong, J.Shen, W.Wang, L. Shao, H. Ling & F. Porikli. (2019). Dynamical hyperparameter optimization via deep reinforcement learning in tracking. IEEE transactions on pattern analysis and machine intelligence, 43(5), 1515-1529. https://ieeexplore.ieee.org/abstract/document/8918068/

[4] S. Pal, A. Pramanik, J. Maiti, & P. Mitra. (2021). Deep learning in multi-object detection and tracking: state of the art. Applied Intelligence, 51, 6400-6429. https://link.springer.com/article/10.1007/s10489-021-02293-7