

# Efficient Siamese Network with Global Correlation for Single-object Tracking

Yufei Lin\*

Department of Robot Science and Engineering, Northeastern University, Shenyang, China

## Introduction

Single-Object Tracking (SOT) is a fundamental problem in computer vision, where the goal is to track the movement of a single object across a sequence of frames in a video. This problem is challenging due to various factors such as object appearance changes, occlusions, and varying lighting conditions. The task becomes even more difficult when tracking objects in real-time or on resource-constrained devices, where computational efficiency is a critical concern. In recent years, Siamese networks have emerged as a powerful framework for SOT due to their ability to learn discriminative features from pairs of images. However, while Siamese networks have shown great promise, challenges remain in improving their efficiency and tracking performance, especially in terms of handling large amounts of data and maintaining robustness in dynamic environments. The proposed approach, a lightweight Siamese network with global correlation for single-object tracking, seeks to address these challenges by focusing on computational efficiency and performance. The key to the success of this approach lies in combining the strengths of Siamese networks with a global correlation mechanism, which enhances the ability to capture long-range dependencies between objects in different frames. By leveraging a lightweight network architecture and global correlation, the model can achieve high tracking accuracy while minimizing computational cost, making it suitable for real-time applications and deployment on devices with limited resources.

## Description

At the core of the Siamese network architecture is the use of twin neural networks that share weights and are trained to compare two images or feature maps. This network structure is well-suited for tracking, as it can compute a similarity score between the target object and candidate regions in the current frame. The Siamese network extracts high-level features from both the template image (the object to be tracked) and the search image (the current frame), and then the network computes a similarity score between the two. This score is used to locate the object in the search image. The simplicity of Siamese networks allows them to be trained effectively with relatively small datasets, making them a powerful tool for tracking. Despite their effectiveness, traditional Siamese networks often suffer from limitations in capturing long-range dependencies between the object and the surrounding context. This is especially problematic when the object undergoes significant motion, rotation, or scale changes, or when it is occluded by other objects in the scene. To address this issue, the proposed model introduces a global correlation mechanism, which allows the network to consider the global context of the object rather than relying solely on local information from the current frame. By computing a global correlation map, the network is better able to understand the relationship between the target object and its environment, thereby

improving the tracking accuracy and robustness to various challenges [1].

The global correlation mechanism works by comparing the features of the template image with the features of the search image at different spatial locations. This process allows the network to compute a correlation map that highlights regions of the search image that are similar to the template. The global correlation map helps the network to focus on the most relevant regions of the frame and disregard irrelevant areas, thereby improving the model's ability to track the object even in the presence of occlusions or significant motion. Additionally, the global correlation mechanism can enhance the network's robustness to scale and viewpoint changes, making it more adaptable to real-world tracking scenarios. To further enhance the efficiency of the model, a lightweight architecture is employed in the proposed Siamese network. Traditional deep learning models, especially those used in image recognition or tracking, can be computationally expensive and require significant processing power, which makes them difficult to deploy on devices with limited resources, such as embedded systems or mobile devices. The lightweight architecture of the proposed model reduces the number of parameters and the computational complexity while maintaining high performance. This is achieved through the use of more compact network layers, such as depthwise separable convolutions, which reduce the computational burden without sacrificing tracking accuracy [2].

In addition to the global correlation mechanism and lightweight architecture, the proposed model also benefits from several other optimizations that enhance its performance. One key optimization is the use of a multi-scale feature extraction approach, which allows the model to track objects at different scales. By extracting features at multiple resolutions, the network can adapt to objects that change in size during the tracking process, improving robustness to scale variations. This is particularly important in real-world scenarios where the object may move closer to or farther from the camera, causing significant changes in its size. Another important optimization is the use of a temporal consistency loss function, which encourages the network to maintain stable object tracking across frames. In SOT, sudden jumps or drifts in the predicted object location can lead to tracking failure. By introducing a temporal consistency loss, the model is guided to make smooth transitions between consecutive frames, reducing the likelihood of such errors. This helps to improve the long-term tracking stability of the network, ensuring that the object is consistently tracked over time, even in challenging conditions such as fast motion or occlusions [3].

The efficiency and effectiveness of the proposed lightweight Siamese network with global correlation are demonstrated through a series of experiments conducted on standard benchmark datasets for single-object tracking, such as OTB-100 and VOT2018. These datasets consist of a diverse set of video sequences with varying challenges, including occlusions, illumination changes, and background clutter. The model's performance is evaluated based on several metrics, including tracking accuracy, robustness, and speed. The results show that the proposed model outperforms existing Siamese network-based trackers in terms of both accuracy and computational efficiency, making it suitable for real-time applications. The use of the lightweight Siamese network with global correlation also has significant implications for the deployment of tracking systems in real-world scenarios. For example, in autonomous vehicles, real-time object tracking is crucial for understanding the surrounding environment and making driving decisions. The lightweight nature of the proposed model ensures that it can be deployed on the limited computational resources available in such vehicles, without compromising tracking performance. Similarly, in security applications, such

\*Address for Correspondence: Yufei Lin, Department of Robot Science and Engineering, Northeastern University, Shenyang, China, E-mail: linyufei@gmail.com

Copyright: © 2024 Lin Y. This is an open-access article distributed under the terms of the creative commons attribution license which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Received: 02 November, 2024, Manuscript No. jtsm-24-157024; Editor Assigned: 04 November, 2024, PreQC No. P-157024; Reviewed: 16 November, 2024, QC No. Q-157024; Revised: 22 November, 2024, Manuscript No. R-157024; Published: 29 November, 2024, DOI: 10.37421/2167-0919.2024.13.472

as surveillance systems, the ability to track objects in real-time with minimal computational overhead is important for maintaining efficiency in large-scale monitoring systems [4].

Another potential application of the proposed model is in robotics, where single-object tracking is essential for tasks such as object manipulation, navigation, and interaction with the environment. In such scenarios, robots must be able to track objects accurately and efficiently while navigating dynamic environments. The global correlation mechanism enhances the model's ability to handle complex environments, while the lightweight architecture ensures that the system can operate on embedded platforms with limited computational resources. The future work in this area could explore further optimizations to improve the performance of the Siamese network, such as the integration of more advanced attention mechanisms, which can help the model focus even more selectively on the most relevant parts of the frame. Additionally, the model could be extended to handle multi-object tracking, which presents additional challenges due to the increased number of objects in the scene and the potential for occlusions and interactions between objects [5].

---

## Conclusion

In conclusion, the lightweight Siamese network with global correlation represents a significant advancement in the field of single-object tracking. By combining the strengths of Siamese networks with a global correlation mechanism, the model is able to capture long-range dependencies and improve tracking accuracy, while the lightweight architecture ensures computational efficiency. The proposed model demonstrates strong performance in various benchmark datasets and has significant potential for real-world applications, particularly in areas such as autonomous driving, security, and robotics. As the field of computer vision continues to evolve, the proposed approach could serve as a foundation for developing even more advanced and efficient tracking systems.

---

## References

1. Huang, Lianghua, Xin Zhao and Kaiqi Huang. "Got-10k: A large high-diversity benchmark for generic object tracking in the wild." *IEEE Trans Pattern Anal Mach Intell* 43 (2019): 1562-1577.
2. Hu, Weiming, Qiang Wang, Li Zhang and Luca Bertinetto, et al. "Siammask: A framework for fast online object tracking and segmentation." *IEEE Trans Pattern Anal Mach Intell* 45 (2023): 3072-3089.
3. Liu, Yun, Zhongsheng Yan, Jinge Tan and Yuche Li. "Multi-purpose oriented single nighttime image haze removal based on unified variational retinex model." *IEEE Trans Circuits Syst Video Technol* 33 (2022): 1643-1657.
4. Hassaballah, Mahmoud, Mourad A. Kenk, Khan Muhammad and Shervin Minaee. "Vehicle detection and tracking in adverse weather using a deep learning framework." *IEEE Trans Intell Transp Syst* 22 (2020): 4230-4242.
5. Russakovsky, Olga, Jia Deng, Hao Su and Jonathan Krause, et al. "Imagenet large scale visual recognition challenge." *Int J Comput Vis* 115 (2015): 211-252.

**How to cite this article:** Lin, Yufei. "Efficient Siamese Network with Global Correlation for Single-object Tracking." *J Telecommun Syst Manage* 13 (2024): 472.