

Selective Information Flow for Transformer Tracking

Janani Kugarajeevan^{a,b}, Kokul Thanikasalam^{c,*}, Amirthalingam Ramanan^c,
Subha Fernando^a

^a*Department of Computational Mathematics, University of Moratuwa, Sri Lanka*

^b*Department of Interdisciplinary Studies, University of Jaffna, Sri Lanka*

^c*Department of Computer Science, University of Jaffna, Sri Lanka*

Abstract

Fully transformer-based one-stream trackers have demonstrated outstanding performance on challenging benchmark datasets over the past three years. These trackers enable bidirectional information flow between all target template and search region tokens to compute attention features without assessing its impact on the tracker’s discriminative capability. Our study found that the information flow from a large portion of background tokens in the search region diminishes the importance of the target-specific features of the template. Moreover, previous transformer-based trackers fail to consider cues from the dynamic background region, even though it contains information about distracting similar objects. To address the identified limitations in transformer tracking, we propose a novel Selective Information Flow Tracking (SIFTrack) framework to enhance the tracker’s discriminative capability by selectively allowing information flow between different types of token groups. In the early encoder layers of the proposed SIFTrack, interactions from all search tokens to target template tokens are blocked to enrich target-specific feature extraction. In the deeper encoder layers, search tokens are partitioned into target and non-target tokens based on their attention scores. Then, bidirectional flow from target search tokens to template tokens is allowed to capture the appearance changes of the target. In addition, by including tokens from the dynamic background, SIFTrack effectively avoids distractor objects by capturing cues from the area surrounding the target. The proposed

*Corresponding author

Email addresses: jananitha@univ.jfn.ac.lk (Janani Kugarajeevan),
kokul@univ.jfn.ac.lk (Kokul Thanikasalam), a.ramanan@univ.jfn.ac.lk
(Amirthalingam Ramanan), subhaf@uom.lk (Subha Fernando)

SIFTrack demonstrated outstanding performance in challenging benchmarks, particularly excelling in the one-shot tracking benchmark GOT-10k, achieving an average overlap of 74.6%. The code, models, and results of this work are available at <https://github.com/JananiKugaa/SIFTrack>.

Keywords: Transformer Tracking, Visual Object Tracking, Vision Transformer, One-stream Tracking

Expert Systems With Applications 259 (2025) 125381