OVERVIEW

**Goal:** enable the tracking of multiple arbitrary objects in real time.

**How it works:** SiamMT is a fully convolutional network that applies individual visual tracking techniques to multiple objects in an efficient and scalable manner. This is accomplished through the global extraction of the frame features, the inclusion of a feature crop-and-resize module and the addition of a new and optimized similarity operation.

CONCLUSIONS

- SiamMT is the first deep-learning-based arbitrary multi-object tracker.
- Tracking quality similar so SiamFC (diff. less than 10%), and surpasses it for some of the multi-object benchmarks.
- Able to track multiple simultaneous objects in real-time, achieving 25 fps with 60 objects for VGA videos and 40 objects for HD720 videos.
- Able to reuse SiamFC’s weights, suggesting the extensibility of SiamMT to other architectures.