EnViD: A Novel Predictive Vision Framework With Temporal Contrast for Resource-Efficient Single Object Tracking in Dynamic Visual Fields

Padmanabhan, Sritej (School: North Allegheny Senior High School)

Autonomous unmanned vehicles (drones, etc.) require fast and resource-efficient perception and control to perform in dynamic, unpredictable real-world environments. Comprehending visual reality requires the efficient acquisition of common-sense knowledge over several regularities in the visual world, e.g., object spatial relationships, temporal dynamics such as motion & rotation, and illuminations on objects in view. For over 40 years, a paramount challenge has resided in the performance of vision systems within dynamic environments such as glare and shadows, particularly in the Single Object Tracking (SOT) domain. Prevailing methods, while sophisticated, are hindered by large datasets and high computational needs with their performance being compromised under varying ambient conditions. Considering these challenges, a novel event-based predictive vision framework (EnViD) was developed to better capture such real-world dynamics. EnViD builds sophisticated and meaningful high-to-high associations, in comparison to high-to-low mappings. This generates semantically rich features, which can be associated with more abstract, low-dimensional concepts through a predictive vision framework. EnViD was tested on an NVIDIA Jetson Nano with real-world conditions such as occlusions and shadows and was shown to surpass current success, precision, and accuracy benchmarks by ~20%. The computational burden of typical frame-based vision was addressed with a temporal contrast signal processing method, similar to the eye's retina, resulting in reduction of data usage by 92%. EnViD's increased efficiency and performance in dynamic conditions make it suitable for many vision tasks. It was proven to improve Telehealth Parkinson's video analysis, with potential use across multiple industries.

Awards Won:
Third Award of $1,000