

Title: An Integrated Computational Model of Visual Search Combining Eccentricity, Bottom-up, and Top-down Cues

Author(s): Gupta, Shashi Kant

Supervisor(s): Venkatesh, K S

Keyword(s): Visual Search, Search Asymmetry, Eye Movement, Deep Neural Networks, Visual Cognition, Computational Modelling

Subject(s): Computer Vision, Deep Learning, Cognitive Science

Abstract: Visual search is described as searching for some target object in a given visual scene, having several other non-target objects. Humans are continuously involved in such tasks in their day-to-day life like searching for a specific food item in the mall or searching for their friends at a party. Extensive studies in visual search behaviour have demonstrated the complex interplay of the target object, the search space, and the humans' memory. In parallel, neurophysiological studies have shown how neuronal circuits form complex visual representations. But very little work has been done that links these exciting works from behavioural studies and neuroscience. This thesis introduces an integrated computational model of visual search that incorporates theoretical frameworks from psychology, resembling the architecture from neurophysiology. The proposed model integrates three essential components, an eccentricity-dependent deep convolutional neural network as a visual processor, top-down target modulated activation maps, and bottom-up saliency-based activations. The proposed model can replicate the standard results from several behavioural studies conducted in visual search literature. And at the same time, it is also efficient enough to search for a target object in a complex natural scene. Various autonomous systems can also significantly benefit from the proposed model, for example, autonomous navigation or visual clinical diagnosis. Most of the previous work on computational modelling of visual search in computer vision involves extensive category-specific training and bear minimal resemblance to biological plausibility. In comparison, the proposed model is self-sufficient and does not require human supervision or extensive task-specific training to search for any new target object. An explainable visual search model that could replicate human visual search behaviour will also bring more trust in such autonomous systems.

