

How is visual search guided by shape? Using features from deep learning to understand preattentive "shape space"

Krista Ehinger, Jeremy Wolfe

Meeting abstract presented at VSS 2016

Visual search can be guided by target shape, but our understanding of how shape guides search has been limited to a set of specific shape features such as curvature, closure, line termination, aspect ratio, or intersection type (see Wolfe & Horowitz, 2004, for a review). These features, while important, do not capture the full range of preattentive shape processing. Understanding how shape guides search more generally requires a model of preattentive "shape space" that correctly represents the similarity between a target shape and different types of distractors. Here, we investigate whether the features learned by "deep learning" convolutional neural networks (CNNs) can be used as a proxy for this shape space. Previous work has shown that the visual representations learned by these networks generalize surprisingly well to a range of visual tasks (Razavian, Azizpour, Sullivan, & Carlsson, 2014). Eight participants performed a visual search task where they searched for a randomly-rotated shape target (a butterfly or rabbit silhouette) among different types of randomly-rotated distractor shapes generated from a family of radial frequency patterns. To characterize the distractor shapes, we ran them through a CNN (Krizhevsky, Sutskever, & Hinton, 2012) and used the feature vector produced by the second-to-last layer of the network as candidate shape features. Easy and hard distractors for each target were well-separated in this shape space, and hard distractors tended to be closer to the target in the neural network's representation of shape. Different participants tended to converge to a similar part of the feature space for hard distractors, but there was less agreement on which distractors were easiest. Our results suggest that the visual representation learned by a "deep learning" CNN is a reasonable approximation of the perceptual space in which humans process shape.

Journal of Vision September 2016, Vol.16, 695. doi:10.1167/16.12.695