Disentangling Generativity in Visual Cognition

Declan Campbell
University of Wisconsin – Madison, Madison, Wisconsin, United States

Timothy Rogers
University of Wisconsin- Madison, Madison, Wisconsin, United States

Abstract

Human knowledge is generative: from everyday learning people extract latent features that can recombine to produce new imagined forms. This ability is critical to cognition, but its computational bases remain elusive. Recent research with -regularized Variational Autoencoders (-VAE) suggests that generativity in visual cognition may depend on learning disentangled (localist) feature representations. We tested this proposal by training -VAEs and standard autoencoders to reconstruct bitmaps showing a single object varying in shape, size, location, and color, and manipulating hyperparameters to produce differentially-entangled feature representations. These models showed variable generativity, with some standard autoencoders capable of near-perfect reconstruction of 43 trillion images after training on just 2000. However, constrained -VAEs were unable to reconstruct images reflecting feature combinations which were systematically withheld during training (e.g. all blue circles). Thus, deep auto-encoders may provide a promising tool for understanding visual generativity and potentially other aspects of visual cognition.