Deep Convolutional Neural Networks Rely on Distinct Semantic Features of Same-Category Objects for Recognition

MohammadHossein NikiMaleki[1], Hamid Karimi-Rouzbahani[2,3]

[1]Faculty of Computer Science and Engineering, Shahid Beheshti University, Iran

[2]Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge, UK

[3]Department of Computing, Macquarie University, Australia

Deep Convolutional Neural Networks (DCNNs) have not only outperformed computer vision algorithms in many applications, but are among the best models of human object recognition. It was recently shown that humans relid on specific segments of objects (called MIRCs) for recognition (Ullman et al., 2016). However, DCNNs did not show such sensitivity to dentical MIRCs. Therefore, it remains unclear if humans and DCNNs use similar strategies for object recognition. One critical question is whether DCNNs rely on semantically similar MIRCs from different exemplars of the same object category. To that end, we obtained MIRCs from one of the most brain-like DCNNs (VGG16), using the well-established Bubbles method (Gosselin and Schyns, 2001). As an advantage to previous precedures, which detected MIRCs from pre-selected discrete image parts, Bubbles sweeps the whole image using continuous masks, allowing data-driven evaluation of all pixels. We obtained MIRCs from 12 categories (each with 16 examplars) of the naturalistic ImageNet dataset (Deng et al., 2009). Results showed clearly different MIRCs for distinct examplars from the same object category, reflecting the highly variable nature of feature extraction in DCNNs, potentially facilitating recognition under image variations. We are collecting human data to quantitatively compare to our DCNN results.

References:

Ullman, S., Assif, L., Fetaya, E. and Harari, D., 2016. Atoms of recognition in human and computer vision. Proceedings of the National Academy of Sciences, 113(10), pp.2744-2749.

Gosselin, F. and Schyns, P.G., 2001. Bubbles: a technique to reveal the use of information in recognition tasks. Vision research, 41(17), pp.2261-2271.

Deng, J., Dong, W., Socher, R., Li, L.J., Li, K. and Fei-Fei, L., 2009, June. Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee.