CONVOLUTIONAL NEURAL NETWORKS TRAINED WITH A DEVELOPMENTAL SEQUENCE OF BLURRY TO CLEAR IMAGES REVEAL DIFFERENT GENERALIZATION CAPABILITIES IN FACE AND OBJECT PROCESSING

Anonymous authors
Paper under double-blind review

ABSTRACT

Convolutional neural networks (CNNs) provide a promising model of human visual processing. However, CNN performance can become severely impaired by modest changes in viewing conditions, such as image blur, whereas human vision is far more robust. What factors might account for the robustness of human vision to blur? Unlike the typical training regime for CNNs, early human visual experience is defined by poor visual acuity at birth followed by gradual improvement over many months. In this study, we examined whether training CNNs with a sequence of blurry to clear images would allow them to achieve robust performance to varying levels of blur. We tested both face and object recognition tasks, two of the most common tasks in daily life. When CNNs were trained on blurry to clear faces, they showed much greater robustness to blur than those trained exclusively on clear faces, consistent with a previous finding by Vogelsang and colleagues (2018). However, when CNNs were trained on blurry to clear objects, they quickly lost whatever initial robustness they had for blurry objects. This difference was clearly revealed by our analyses of spatial frequency tuning preferences of the networks. The blurry to clear face-trained network maintained a preference for lower spatial frequencies throughout training, whereas the blurry to clear object-trained network showed a progressive shift in tuning preferences towards higher spatial frequencies. Control studies performed with objects, matched to the Fourier amplitude spectrum of the faces, indicated the pervasiveness of these effects. Our findings may reveal a fundamental difference in face and object recognition, such that faces can be processed in a more holistic manner. Although a progression of blurry to clear faces in infancy might be beneficial for face processing, our findings suggest that such early experiences with objects are insufficient to account for human robustness to object blur.