Automated delineation of visual area boundaries and eccentricities by a CNN using functional, anatomical, and diffusion-weighted MRI data

Noah C. Benson[1], Bogeng Song[2], Toshikazu Miyata[3,4,5], Hiromasa Takemura[3,4,5,6], Jonathan Winawer[2,7]

[1] eScience Institute, University of Washington; [2] Dept. of Psychology, New York University; [3] Graduate School of Frontier Biosciences, Osaka University; [4] Division of Sensory and Cognitive Brain Mapping, Department of System Neuroscience, National Institute for Physiological Sciences; [5] Center for Information and Neural Networks (CiNet), Advanced ICT Research Institute, National Institute of Information and Communications Technology (NICT); [6] Department of Physiological Sciences, School of Life Science, SOKENDAI (The Graduate University for Advanced Studies); [7] Center for Neural Sciences, New York University

Abstract. Delineating visual field maps and iso-eccentricities from fMRI data is an important but time-consuming task for many neuroimaging studies on the human visual cortex because the traditional methods of doing so using retinotopic mapping experiments require substantial expertise as well as scanner, computer, and human time. Automated methods based on gray-matter anatomy or a combination of anatomy and functional mapping can reduce these requirements but are less accurate than experts. Convolutional Neural Networks (CNNs) are powerful tools for automated medical image segmentation. We hypothesize that CNNs can define visual area boundaries with high accuracy. We trained U-Net CNNs with ResNet18 backbones to predict either V1, V2, and V3 boundaries or 5 regions of iso-eccentricity using human-labeled maps. Separate CNNs were trained to predict these regions using different combinations of the following input data: (1) anatomical data from a T1-weighted image only, (2) anatomical data from T1-weighted and T2*-weighted images, (3) white-matter tract endpoints from diffusion-weighted imaging, (4) functional data from retinotopic mapping. All CNNs using functional data had cross-validated accuracy that was statistically indistinguishable from the inter-rater reliability of the training dataset (dice coefficient of 92%) while the CNNs lacking functional data had lower but similar accuracies (~75%). Existing models that do not use CNNs had accuracies lower than any of the CNNs. These results demonstrate that with current methods and data quality, CNNs can replace the time and effort of human experts in manually defining early retinotopic maps, but cannot yet replace the acquisition of functional data.