How many non-linear computations are required for CNNs to account for the response properties of the primary visual cortex (V1)?

Hui-Yuan Miao, Hojin Jang, and Frank Tong
Department of Psychology, Vanderbilt University

Introduction

Background
• The primary visual cortex (V1) has been extensively studied but the computations performed by V1 neurons are not yet fully understood.
• Prominent V1 models include Gabor-based models as well as simple linear-nonlinear models with learned filters.

Example of Gabor-based V1 model

However, a recent study found that unit responses in an intermediate layer of a deep convolutional neural network (VGG-19) outperformed the best traditional V1 models at predicting neuronal responses to natural images, suggesting that V1 neurons perform far more non-linear computations than previously expected.

Potential concerns
• VGG-19 uses small convolutional filters whereas the image inputs used in Cadena et al. (2019) were comparatively large. Thus, we suspected that the poor performance of the lower layers of VGG-19 may have been driven by the mismatch between convolutional filter size and image input size.

Methods and computational modeling

Dataset
The dataset from Cadena et al. (2019) consists of recordings of 166 neurons from 2 macaque monkeys while viewing 7250 natural or synthetic images.

Example stimuli
1450 ImageNet images
5000 synthetic images made from 1450 ImageNet images

Example of simple linear-nonlinear model with learned filters

Results: AlexNet vs. VGG-19

Matching VGG-19 and AlexNet performance
Given that VGG-19 still attains higher performance, can a modified version of AlexNet (with an additional convolutional layer before pooling and more channels in conv2) match the predictive performance of conv3_1 in VGG-19?

Testing non-linear transformations with the Gabor-based V1 model
All nonlinear transformations led to small but statistically significant improvements of the Gabor-based V1 model.

Conclusions
While V1 is not purely linear, the number of non-linear computations needed to explain V1 neuronal responses is far less than recently claimed by Cadena et al. However, CNN-based models outperformed the Gabor-based model, suggesting that CNNs can provide a new state-of-the-art V1 model.

References