**V1 Model**

For every image, the V1 model produces the responses of a population of putative V1 neurons using a bank of Gabor filters with lateral interactions that implement divisive normalization. In the V1 model, we used a scaling factor of 1 dva = 18.2 pixels based on published observations regarding the sizes of V1 receptive fields. We rescaled the images in our set to 85 pixels (along the longer dimension) so that the image size seen by the V1 model matched the image size used in our experiments. The V1 model is identical to a previously published model (Pinto et al., 2008) and is summarized below. V1 responses were obtained in four steps.

1. Each image was first normalized for intensity and contrast by subtracting its mean and dividing by its standard deviation.
2. This normalized image was subjected to divisive normalization: for each 3 x 3 pixel window, we subtracted the mean and divided each pixel by the Euclidean norm of the pixels in the window. This normalization was constrained to reduce but not enhance responses (i.e. we divided only if the norm was greater than 1.
3. The resulting image was given as input to 96 V1 model neurons each having a linear Gabor kernel. Each filter had a size of 43 x 43 pixels with a standard deviation of 9 cycles/pixel. The filters were tuned to 16 equally spaced orientations from 0 to 360 degrees, with 6 spatial frequencies (2, 3, 4, 6, 11 and 18 pixels per cycle). We normalized the output of each of the 96 Gabor filters to have zero mean and a norm of 1. The output of each Gabor filter was then passed through a response nonlinearity comprising a threshold function (all negative values were set to zero), and response saturation (all values greater than 1 were set to 1).
4. Finally, we performed output divisive normalization, where for each filter output, we subtracted the mean of filter outputs in a fixed spatial window of 3 x 3 pixels across all orientations and spatial frequencies and then divided by the Euclidean norm of all values in the spatial window of 864 elements (3 x 3 x 96), if the norm was greater than 1.

The output of the V1 model consisted of the activity of 96 neurons in response to an image, which was taken to be a vector in 96-dimensional space.