Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/7/1/eabd6127/DC1

Supplementary Materials for

Visual number sense in untrained deep neural networks

Gwangsu Kim, Jaeson Jang, Seungdae Baek, Min Song, Se-Bum Paik*

*Corresponding author. Email: sbpaik@kaist.ac.kr

Published 1 January 2021, *Sci. Adv.* **7**, eabd6127 (2021)
DOI: 10.1126/sciadv.abd6127

This PDF file includes:

Figs. S1 to S4
Table S1

Supplementary Materials

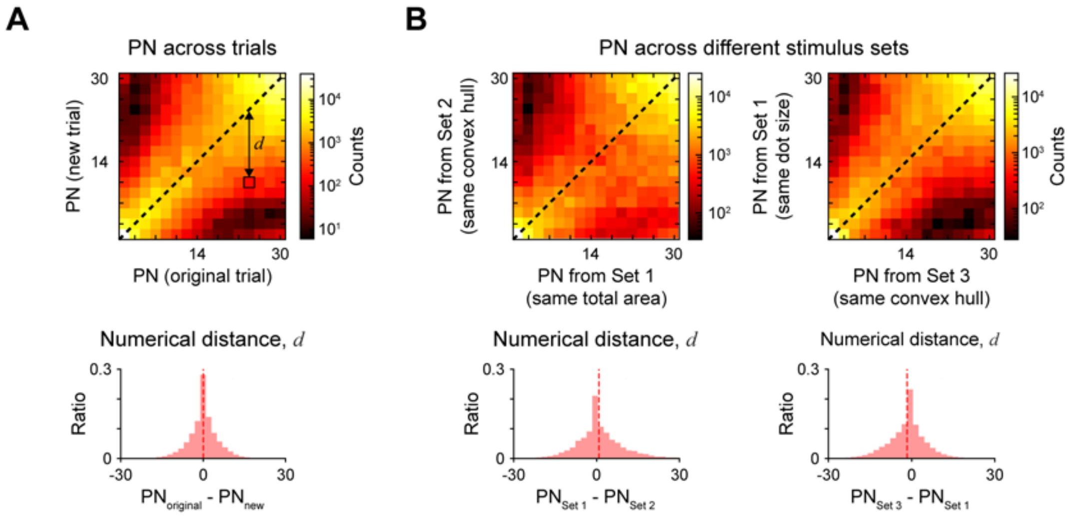

Fig. S1. Consistency of the preferred numerosity of number neurons.

(**A-B**) Top: The preferred numerosity (PN) measured with new stimulus sets of the same type (**A**) or with different stimulus types (**B**) are significantly correlated with each other, implying consistency of the preferred numerosity. Bottom: The average numerical distance between preferred numerosities of each number-selective neuron measured with different stimulus conditions is close to zero. Dashed lines indicate the average.

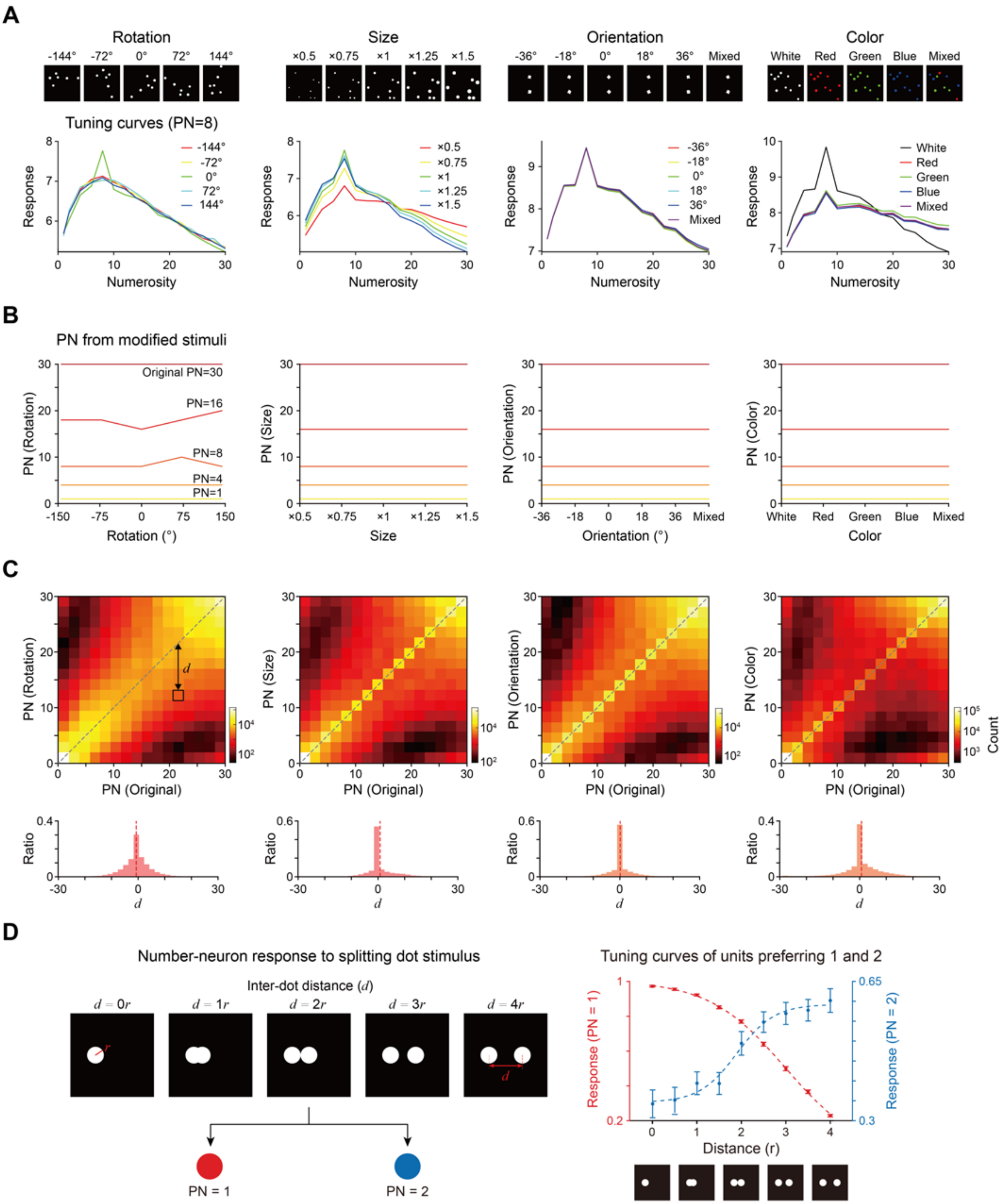

Fig. S2. Number tuning independent of other low-level visual cues.

(**A**) Top: Additional analyses were performed to confirm that the observed preferred numerosity remains consistent under (1) rotation of images, (2) variation of the dot size, (3) rotation of each (square) dot, and (4) dot color. Bottom: Sample tuning curves of PN = 8 units are shown for various stimulus conditions. **(B)** The response of number-selective neurons to the preferred numerosity (peak value of the average tuning curve) remained consistently stronger than that to other numerosities, under the variation of low-level visual cues of the stimulus. Observed PN of the sample units (PN = 1, 4, 8, 16, 30) are shown. (**C**) Top: The preferred numerosity (PN) outcomes measured with the original and modified stimulus conditions are significantly correlated with each other. Bottom: The average numerical distance between PNs measured with different stimulus conditions is close to zero. (**D**) Responses of neurons preferring the numbers 1 and 2 as the distance between two stimulus dots vary. Left: Design of the stimuli. Two identically sized dots (total area – 226 pixel^2) were located at random locations (50 iterations) and one dot was gradually shifted toward the other dot until the two dots fully overlapped. Right: The responses of number-selective neurons preferring 1 or 2 to these stimulus images (units of top 30% sharp tuning width in logarithmic number scale). For each initialization condition, the average responses of PN = 1 and PN = 2 units were normalized from 0 to 1. The response of each neuron was greater when their preferred numerosity (1 or 2) was given regardless of the distance between the two dots. The response of number-selective neurons preferring 1 remained fairly high until the elongated dot is split into two dots (distance at approximately 2r), and then the response noticeably decreased as the two split dots move further apart. In contrast, the response of number-selective neurons preferring 2 significantly increased when the elongated dot was split into two parts and remains consistent as the distance between the two dots increased. Dashed lines indicate fitted logistic curves, PN1: R^2 = 0.9997, midpoint = 2.98, PN2: R^2 = 0.9844, midpoint = 1.90.

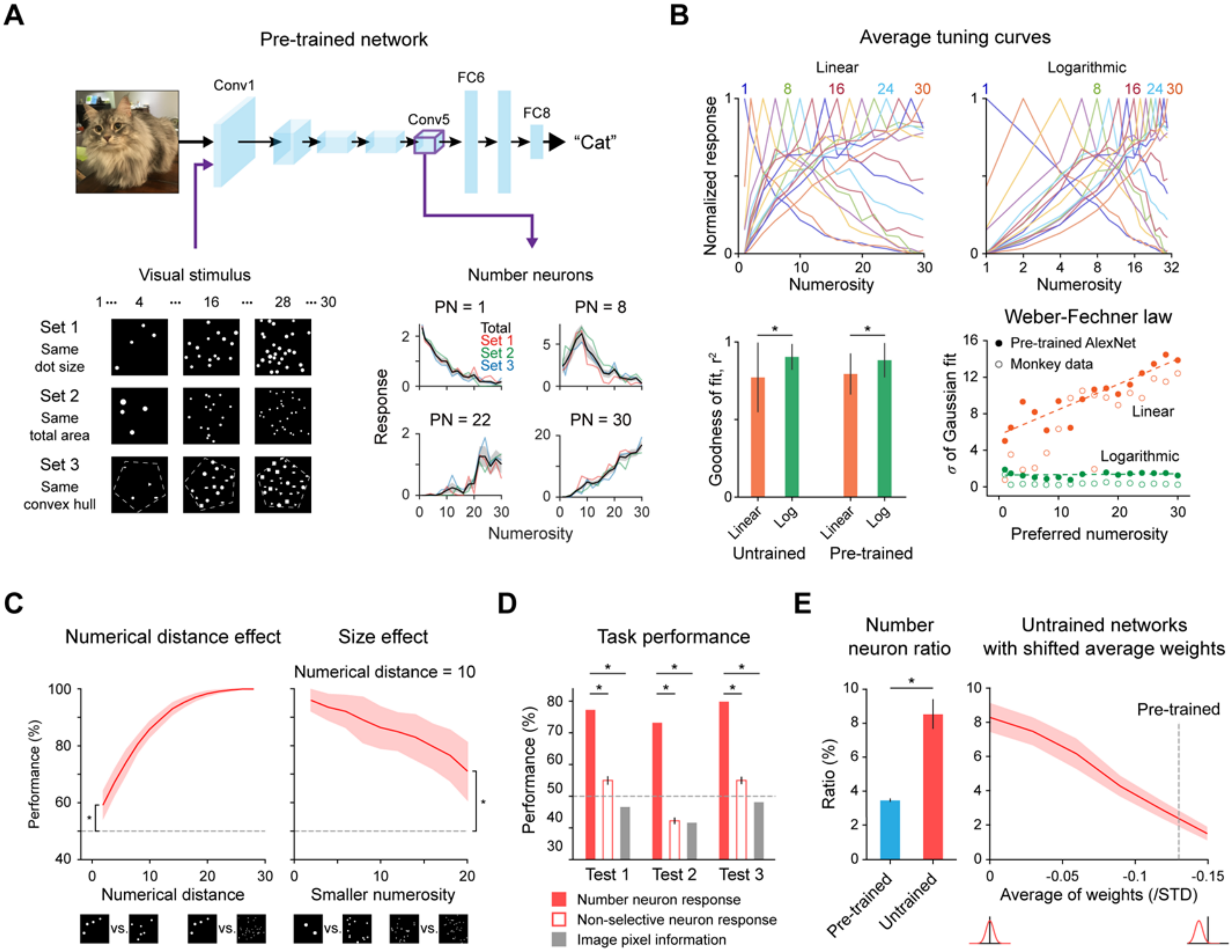

Fig. S3. Number-selective neurons in pre-trained AlexNet.

(**A**) Top: Architecture of the pre-trained AlexNet (Photo credit: Se-Bum Paik, KAIST). Left bottom: Examples of the stimuli used to measure number tuning, as shown in Fig. 1A. Right bottom: Examples of tuning curves of individual number-selective neurons as reported (*21*). **(B-D)** Most tuning properties observed in untrained networks, including the Weber-Fechner law (Fig. 2C) and the numerical distance/size effects (Fig. 3D), were also reproduced in pre-trained networks. **(B)** Average tuning curves, **(C)** Distance and size effect, and **(D)** Numerosity comparison task performance of number-selective neurons in the pre-trained AlexNet. (**E**) Left: The ratio of number-selective neurons to the total neurons is significantly smaller in the pre-trained network ($P < 10^{-40}$, Wilcoxon rank-sum test). Right: The ratio of number-selective neurons is dependent on the bias of the average convolutional weights in the pre-trained network. When the weights of untrained networks are shifted with a negative bias, the ratio of number-selective neurons decreases similar to that of pre-trained network.

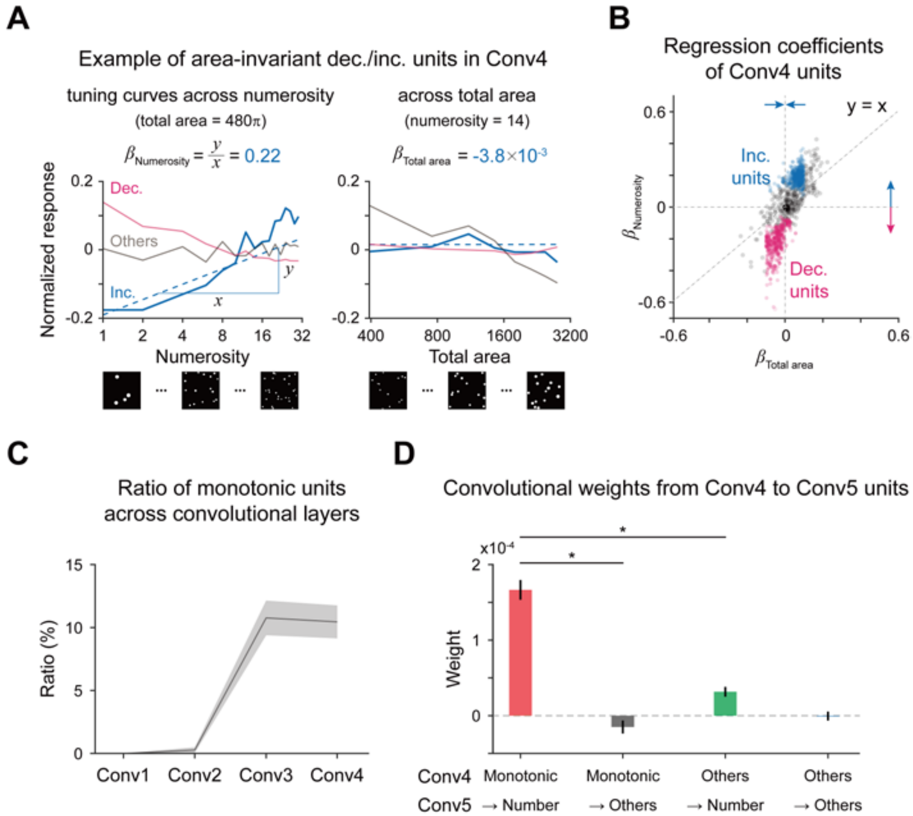

Fig. S4. Stimulus area-invariant, monotonically increasing and decreasing unit activities.

(**A**) Examples of tuning curves of increasing (blue), decreasing (pink), and other (gray solid line) network units in Conv4 of the untrained AlexNet are shown. As suggested in previous work (*16*), increasing and decreasing units were defined when the response monotonically changes as the numerosity increases (left), but remains consistent as the total area of the stimulus increases (right; see Methods for details). Blue dashed lines indicate the regression for defining increasing units. (**B**) Regression coefficient (β) for log (numerosity) and log (cumulative area) of Conv4 units. Pink, blue, and gray dots indicate 500 randomly selected decreasing, increasing, and other units, respectively. (**C**) The ratio of monotonic units is less than 0.3% in Conv1 and Conv2, but it increases up to 10% in Conv3, suggesting that the hierarchical convolutional layers are required to generate monotonic activities. (**D**) Monotonic units in Conv4 provide stronger inputs to number neurons than to the other neurons in Conv5 (red vs. gray; *P = 9.61×10^{-21}, Wilcoxon rank-sum test), implying that number tuning in Conv5 arises from the monotonic units in Conv4. Inversely, number neurons in Conv5 also connect to monotonic units more strongly than the other Conv4 units (red vs. green; *$P = 5.59\times10^{-16}$).

Layer	Type	Number of neurons	Kernels	Activations
Input	Image input	227 × 227 × 3	Weights 11 × 11 × 3 × 96 Bias 1 × 1 × 96	
Conv1	Convolution	55 × 55 × 96		ReLU and cross channel normalization
Pool1	Max pooling	27 × 27 × 96		
Conv2	Convolution	27 × 27× 256	Weights 5 × 5 × 48 × 256 Bias 1 × 1 × 256	ReLU and cross channel normalization
Pool2	Max pooling	13 × 13 × 256		
Conv3	Convolution	13 × 13 × 384	Weights 3 × 3 × 256 × 384 Bias 1 × 1 × 384	ReLU
Conv4	Convolution	13 × 13 × 384	Weights 3 × 3 × 192 × 384 Bias 1 × 1 × 384	ReLU
Conv5	Convolution	13 × 13 × 256	Weights 3 × 3 × 192 × 256 Bias 1 × 1 × 256	ReLU
Pool5	Max pooling	6 × 6 × 256		
FC6	Fully Connected	1 × 1 × 4096	Weights 4096 × 9216 Bias 4096 × 1	ReLU and dropout
FC7	Fully Connected	1 × 1 × 4096	Weights 4096 × 4096 Bias 4096 × 1	ReLU and dropout
FC8	Fully Connected	1 × 1 × 1000	Weights 1000 × 4096 Bias 1000 × 1	Softmax
Output	Classification Output			

Table S1. Summary of the architecture of AlexNet.

The network consists of five convolutional layers for feature extraction (Conv1 – Conv5) and three fully connected layers for object classification (FC6 – FC8). In the current study, to investigate the selective responses of neurons rather than the trained performance of the system, the classification layers were discarded and the responses of units in the last convolutional layer (Conv5) were examined.