

# Relationship between training database size and structural connectivity in artificial neural networks: Do specific local connections improve prediction accuracy?

Min Sik Park (✉ [mspark91@gmail.com](mailto:mspark91@gmail.com))

Samsung Advanced Institute of Technology <https://orcid.org/0000-0003-3466-9979>

---

## Research Article

**Keywords:** Brain-like Neural Network, Bypass connection, Disconnection, Inhibitory nodes, Training database size

**Posted Date:** November 12th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1070921/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# **Relationship between training database size and structural connectivity in artificial neural networks: Do specific local connections improve prediction accuracy?**

Min Sik Park\*

Computer-Aided Engineering Group, Samsung Advanced Institute of Technology, Samsung Electronics, 130 Samsung-ro, Suwon, Gyeonggi-do 16678, Republic of Korea

\* E-mail address: [ms91.park@samsung.com](mailto:ms91.park@samsung.com)

## **Abstract**

Training of artificial neural networks is very expensive, as a large-size database is necessary. Moreover, it is usually difficult to find such large-size training databases. Hence, it will be interesting to design artificial neural networks that can be used for training with a small-size database, while maintaining a similar accuracy for prediction compared to fully connected neural networks. We studied neural networks with partial disconnections, additional bypass connections, and negative activation nodes, which are found in the neuronal systems of the human brain. By combining the fully connected neural network and the above three brain-like elements, we found that the modified neural network showed improved prediction accuracy of 13% compared to the fully connected one despite the small size of the training database. To analyze the improved neural network, the contribution of each node in the hidden layers affecting the total prediction accuracy of the neural networks was studied. We also found important local connections that improve the prediction accuracy, and discussed the design of a neural network with a small-size training database without reduction in prediction accuracy.

*Keywords: Brain-like Neural Network, Bypass connection, Disconnection, Inhibitory nodes,  
Training database size*

## 1. Introduction

The realization of intelligence systems mainly depends on the size of the training database, that is, the larger the size of the training database, the higher the prediction accuracy. However, unlike intelligent machines, human beings can learn faster and from fewer examples. This may be related to the specific structure of neural networks in the human brain. The formation of specific brain structure may have been affected due to evolution, for adaptation to harsh environments, such as fear, starvation, and cold [1], given the need to make decisions quickly. The structure of a neural network can be divided into the physical geometry of neuronal connections (structural connectivity) and the chemical connection strength between the neurons (evoked functional connectivity). Thus, although two brains may have the same physical geometry in terms of neuronal connections, the shape of the flow of its neuronal signals may vary depending on the synaptic strength.

Human brains show distinguishable spatial distributions of functional areas such as the visual, auditory, and prefrontal cortices. It is probable that the connections between neurons are governed by several rules, which may affect the formation of each functional area. However, the connection rules for the functional areas in the human brain have not been revealed yet. According to Hebb's rule, neurons that fire together, wire together [2]. Hebb's rule explains how associative memory can be formed by changing the strength of the synapses. However, Hebb's rule works within a given neuronal structure, that is, it does not explain how the synapses are physically formed between the neurons. It is still unclear whether the synapses between neurons are formed randomly or genetically [3-8]. Then, how can neurons located far apart make connections to create associative memory? In this case, the direct connection between two memories is impossible physically. Instead, they are probably connected indirectly through the higher layers in the hierarchical structure.

Inhibitory neurons exist in addition to excitatory neurons and account for 20–30% of all neurons. The role of inhibitory neurons is not clear yet. However, it is expected that they can

suppress a surplus of neuronal signals and prohibit interference between these signals. In the visual cortex, the role of inhibitory neurons is related to orientation selectivity [9-11]. However, the principles behind the formation of excitatory–inhibitory networks have not been studied as diligently as compared to functional studies of the role of inhibitory neurons. Thus, it is interesting to probe why a neural network composed only of excitatory neurons cannot be optimized without inhibitory neurons.

A fully connected artificial neural network (FNN), a machine learning method, is considered to be a candidate for realizing artificial intelligence. Notably, the emergence of deep learning with big data has accelerated developments of artificial intelligence technology such as image classification, voice recognition, and self-driving cars [12-15]. FNNs use the strength of the connections between nodes (i.e., weights) in the same manner as the control of the strength of synapses in the brain for learning. However, only the numbers of hidden layers and nodes are set up, and the types of connections and nodes have not been considered much in studies on FNNs.

Here, we assumed that the functional connectivity, i.e., trained weight structure, can be affected by the structural connectivity. Therefore, we adopted three properties of neural networks within the human brain, which are not included in the FNNs and well observed in the neuronal network of human brains [16, 17]. The first property is that all nodes in the hidden layers do not need to connect fully between adjacent layers. The second property indicates that additional bypass connections in the hidden layers can exist between the following adjacent layers or beyond. The third property relates to the nodes in the hidden layers, which can be inhibitory instead of excitatory (i.e., negative activation nodes; Fig. 1). In the Eq. 1, the main variables affecting prediction accuracy in FNNs and brain-like neural networks (BNNs) including the above three properties are summarized. Even though these three properties do not represent the whole set of neuronal connections in the brain, they can be used as starting points to study how the brain can learn from small-size training data sets. Therefore, our goal in this work is to reveal

the relationship between the size of the training database and the structural connectivity of BNNs (Fig. 1).

$$\text{Prediction Accuracy} \sim f_{\text{FNN}}(N_{\text{layers}}, N_{\text{nodes}}) \text{ for FNN,} \quad (1)$$

$$\sim f_{\text{BNN}}(N_{\text{layers}}, N_{\text{nodes}}, N_{\text{discon}}, N_{\text{addition}}, N_{\text{inh}}) \text{ for BNN,} \quad (2)$$

where the  $N_{\text{layers}}$ ,  $N_{\text{nodes}}$ ,  $N_{\text{discon}}$ ,  $N_{\text{addition}}$ , and  $N_{\text{inh}}$  correspond to the number of hidden layers, nodes in hidden layers, disconnections, additional bypass connections, and inhibitory nodes, respectively.

## 2. Methodology

**2.1. Fully connected neural networks.** For the optimization of FNNs, 3k FNNs were generated depending on a combination of MNIST database size (6 steps of 10k, 20k, ..., 60k), number of hidden layers (5 steps of 1, 2, ..., 5), number of nodes in each hidden layer (10 steps of 100, 200, ..., 1k), and training epochs (10 steps of 100, 200, ..., 1k). Additionally, 360 FNNs were also generated to study the shape dependence of the hidden layers (6 cases of increase (50-100-150-200-250) / decrease (250-200-150-100-50) / random (200-50-100-250-150) / zigzag-edge (250-50-250-50-250) / zigzag-center (50-250-50-250-50) / maximum (250-250-250-250-250), where each case has 60 steps depending on the database size and epochs). All FNNs used sigmoid activation with bias and a backpropagation learning algorithm was employed with L2 regularization (0.1), adaptive learning rate (0.001), inertia term (momentum, 0.001), mini-batch, and cross-entropy loss functions. We found the gradually decreasing shape of the hidden layers to be the most efficient among all 6 cases. Compared to the maximum shape with the maximum number of weights, the FNNs with the decreasing shape of the hidden layers (250-200-150-100-50) showed similar test accuracy. Therefore, we only considered FNNs and BNNs with a gradually decreasing shape of hidden layers for comparison in the next section.

**2.2. Brain-like neural networks.** The three properties of neuronal networks were applied to the FNNs (Fig. 1). For the disconnection, we selected connections between nodes in adjacent hidden layers to be removed randomly at a given ratio between 0 and 0.5. For additional bypass connections between the following adjacent hidden layers, we defined the bypass level to select target hidden layers to connect. For instance, level 2 referred to additional bypass connections between 250 and 150, 200 and 100, and 150 and 50 in the 250-200-150-100-50 hidden layers (Fig. S1(a)). Level 3 indicated additional bypass connections between 250 and 100, and 200 and 50 (Fig. S1(b)). Similarly, level 4 indicated additional bypass connections between 250 and 50 (Fig. S1(c)). As in the disconnection, a ratio between 0 and 0.5 was assumed to select hidden nodes for

the additional bypass connections randomly. For the inhibitory nodes, the signs of the activation values were modified to become negative, and a ratio between 0 and 0.5 was also applied for the random selection of inhibitory nodes. To find optimized BNNs showing high prediction accuracy, we set up 576 BNNs in a combination of ratios between 0 and 0.5 for the disconnections, additional bypass connections, and inhibitory nodes, and the range of between 2 and 4 for the level in the additional bypass connections (Table I). Other learning conditions, such as backpropagation, sigmoid activation, etc., were the same as those for the FNNs.

**2.3. Analysis method.** The knockout method has been used for the identification of the role of specific neurons in experimental neuroscience [18,19]. Here, the knockout method was developed to quantitatively analyze the contribution of each node to the prediction accuracy in BNNs. After training the BNNs, all connections from a targeted hidden node were removed (i.e., the node was knocked out), and the test accuracy was recalculated.

**2.4. MNIST database.** The MNIST database is a collection of handwritten digits from 0 to 9 [20]. It consists of 60k training and 10k test datasets. Each image in grayscale has a dimension of  $28 \times 28$  pixels and, therefore, the input vectors for FNNs and BNNs had a length of 784. The dimension of the output was 10 (0–9). For training, two types of BNNs were used, consisting of the hidden layers 250-200-150-100-50 and 100-80-60-40-20 for the test and analysis, respectively. Other conditions were the same as those for the other cases using different databases.

**2.5. FEI database.** The FEI database is a collection of face images [21]. It consists of 2,800 images obtained from 14 rotational directions of 200 individuals. Each color image has a dimension of  $640 \times 480$  pixels. Since the FEI DB has a large vector length of 307,200, the size and color of the images were reduced to  $40 \times 30$  (resulting in a length of 1,200) and converted to



grayscale, respectively. The output dimension depends on the number of individuals used for training and the test (1–200). The 13 images for the rotational faces were used for training, and 1 image for a face in a specific rotation was used for the test. For the training, the BNNs consisted of the hidden layers 4000-2000-1000-500-250. Other conditions were the same as those for the other cases using different databases.

**2.6. PubChem database.** The PubChem database is a collection of information on chemical molecules [22]. It consists of 93.9 million compounds, with a simplified molecular-input line-entry system (SMILES). Eighty-six parameters, including 12 atomic elements, 2 bonding types, 22 ring structures with heteroatoms, and 50 functional groups, were extracted from SMILES as molecular descriptors for input vectors of the FNNs and BNNs [23]. The 20k training dataset was selected from CID 1–34,050 organic compounds, and the 56,040 test dataset was also chosen from CID 34,051–136,200 compounds. The outputs are oxidation potentials for each compound. The oxidation potentials of the training dataset were obtained from high-throughput quantum chemistry calculations [23]. For the training, the BNNs consisted of hidden layers 250-200-150-100-50. Other conditions were the same as those for the other cases using different databases.

### 3. Results and discussion

To study the optimized BNN structure, we initially set up the FNNs with five hidden layers. The shape of the hidden layers was optimized by screening 3,480 FNNs using the MNIST database. As per the optimized shape, the first hidden layer next to the input layer should be the largest among the hidden layers, and the sizes of the following hidden layers should decrease gradually.

Three BNN factors, disconnection, additional bypass connection, and inhibitory nodes, were included in the FNNs with the optimized structure of hidden layers. Based on the combinations of the three BNN factors, 576 BNN structures with 250-200-150-100-50 hidden layers were generated and tested for predictions using the MNIST database (Table I). Here, the 20k database was used to test the BNNs for the case of a large-size database training.

Fig. 2(a) shows the test results for 576 combinatorial BNNs considering the increase in the number of inhibitory nodes. Overall, the prediction accuracy gradually decreases as the number of inhibitory nodes increases and, therefore, 6 clusters of points appear distinctly. In addition, the distribution of accuracy depending on the disconnections and additional bypass connections becomes wider as the number of inhibitory nodes increases. However, the order in the prediction accuracy is not conserved in the BNNs with the change in the number of inhibitory nodes. For instance, the 61<sup>st</sup> BNN (depicted by the red circle on the left) shows minimum accuracy (95.71%) among the 96 BNNs without inhibitory nodes (the 1<sup>st</sup> cluster). However, the 169<sup>th</sup> BNN (depicted by the red circle to the right), which has the same structure as the 61<sup>st</sup> BNN, shows higher accuracy (96.14%) as well as a relatively higher order among the 96 BNNs with 10% inhibitory nodes (the 2<sup>nd</sup> cluster). At the beginning of this study, we expected that the minus sign activation of the inhibitory node would not affect the training results since the product of the minus activation and weights can be of the same sign compared to the multiplication between the plus sign activation and weights during training. However, the results show that the inhibitory nodes can affect the performance of the BNNs.

Fig. 2(b) shows the test results for 576 combinatorial BNNs in view of the increase in the ratio of disconnection. The tendency of prediction accuracy shows little change when the disconnection ratio was modified. BNNs showing better prediction accuracy exist even when the range of disconnection is 50%. Fig. 2(c) shows the contribution of additional bypass connections in the 576 BNNs. Compared to the results for the disconnection, the prediction accuracy increases slightly, that is, the negative slopes of 6 clusters become gentler as the ratio of additional bypass connections increases. Fig. 2(d) shows the influence of the level of additional connections from 2 to 4, as indicated by three big clusters in the graph. The dependence of additional bypass connection levels is not clear for the prediction accuracy. On the contrary, the contribution of the inhibitory nodes is most evident in form of the obvious negative slope in each cluster following additional bypass connection levels. The next smaller contribution is apparent in the small 6 clusters pertaining to the ratio of additional bypass connections inside the larger 3 clusters. Notably, the big cluster corresponding to level 3 in the center shows increasing prediction accuracy with the increasing ratio of additional bypass connections in the smaller rate of inhibitory nodes (red dotted circle). The next cluster is the big cluster on the left, corresponding to level 2, which shows increasing prediction accuracy for the increasing ratio of additional bypass connections in the higher ratio of inhibitory nodes. However, the 1<sup>st</sup> and 2<sup>nd</sup> smaller clusters within the big cluster corresponding to level 2 on the left-hand side, and the big cluster of level 4 on the right-hand side, show a weak relationship between additional bypass connections in the BNN structures and prediction accuracy. Hence, it is expected the additional bypass connections related to level 3 will improve the prediction accuracy in the BNNs with 5 hidden layers, compared to the FNN.

In Fig. 3, the role of inhibitory nodes is studied in more detail. Here the activation thresholds of inhibitory nodes are changed from 0 to 100 (0, 1, 5, 100). The increased activation thresholds mean that the inhibitory nodes activate higher inputs compared to the excitatory (positive) nodes. In other words, as the threshold increases, the number of inactive inhibitory nodes increases. Fig.

3 shows the decrease in prediction accuracy as the activation threshold increases for a high ratio of inhibitory nodes. Comparing the thresholds of 5 and 100 in Fig. 3(c) and 3(d), the decrease in prediction accuracy becomes saturated after 5. This result may be related to the fact that the sum of the size of inputs from the front does not exceed 5; instead, the decrease in prediction accuracy may originate from the non-activated inhibitory nodes, similar to the fully disconnected nodes. Therefore, the prediction accuracy decreases further compared to the disconnection cases consisting of partially disconnected nodes. The wider distribution of the prediction accuracy in cases with a higher ratio of inhibitory nodes for all threshold cases indicates the importance of inhibitory node distribution in the BNNs. The role of inhibitory nodes may be related to the multiple database training in the same neural network. If we assume a neural network with different input and output layer nodes and the same hidden layer nodes, we can probably design inhibitory node distribution in a neural network for different accuracies depending on the type of database.

To analyze the BNN showing higher prediction accuracy compared to the FNN, we set up the 576 BNNs based on 100-80-60-40-20 hidden layers and maintained the other parameters as before. Moreover, the small size of the MNIST database (1k) was used for training the BNNs to discover the relation between the local connections of the BNNs and the size of the database. Fig. 4(a) shows the different prediction accuracy tendencies compared to the BNNs using a 20k training database. It is remarkable that the many BNNs with inhibitory nodes show higher prediction accuracy irrespective of their ratio, compared to the FNN with a prediction accuracy of 69.81%. For the analysis by the knockout method, we selected the BNN with 0.2 as the ratio of disconnections for the connections between all hidden layers, 0.5 as the ratio of additional bypass connections for level 3 ( $100 \rightarrow 40$ ,  $80 \rightarrow 20$ ), and 0.5 as the ratio of inhibitory nodes for all hidden layers. The prediction accuracy is 77.48%. The total number of weights used in this BNN is fewer than that in the FNN, as many as 400 weights. Fig. 4(b) represents the analysis results in terms of a bar graph. Each bar graph corresponds to the knockout results of nodes in each hidden

layer. Notable, the additional bypass connections from the 1<sup>st</sup> (100  $\rightarrow$  40) and 2<sup>nd</sup> (80  $\rightarrow$  20) hidden layers are shown in separate bar graphs. In each graph, the contribution of the target nodes is inversely proportional to the height of the bar; that is, a higher contribution of a knockout node induces lower prediction accuracy of the BNN. For the 1<sup>st</sup> (100  $\rightarrow$  80) and 4<sup>th</sup> (40  $\rightarrow$  20) hidden layers, the contributions of each node for the prediction accuracy are larger than that of each node in the 2<sup>nd</sup> (80  $\rightarrow$  60) and 3<sup>rd</sup> (60  $\rightarrow$  40) hidden layers. On the contrary, the additional bypass connections from the 2<sup>nd</sup> (80  $\rightarrow$  20) hidden layer show large contributions to the prediction accuracy compared to those from the 1<sup>st</sup> (100  $\rightarrow$  40) hidden layer. The analysis results obtained from the knockout method are summarized in the center of Fig. 4(b), and the important connections are indicated by curved arrows. Therefore, it is expected that the existence of bypassing paths (additional connections) makes neural networks adjustable for the various sizes of training databases; that is, these paths protect the overfitting of neural networks from the imbalance between the number of weights and the size of the database. When the size of the database is small, the additional connections, instead of their normal counterparts, are used through the bypass. Therefore, if we can reduce the number of weights in the FNN and add the bypassing weights, the adjustability of neural networks for the various sizes of databases will be enhanced for a similar computational cost.

To study the database dependency of the BNNs, we used three kinds of databases: MNIST, FEI, and PubChem [8-10]. Since these databases are well studied in the range of vision and material science, they are appropriate for the test of BNNs. The characteristics of the three databases are listed in Table II. The most evident aspect is that the size of the MNIST and PubChem databases is large, whereas that of the FEI database is small. Moreover, the MNIST and FEI databases consist of images for digits and faces, respectively, whereas the PubChem database refers to molecular properties, that is, it contains descriptors. Therefore, we can study the BNNs depending on the size and character of three different types of databases. Fig. 5 shows the prediction accuracy for the 576 BNN structures generated by combinations of three BNN factors. For the

MNIST and PubChem databases, the prediction accuracy is not affected much by the BNN structure. On the contrary, the prediction result for the FEI database shows a large dependence on the BNN structures. The size of the training FEI database (= 390) is very small compared to that of the other two databases (= 20k). An improvement of up to 13% occurs in the prediction accuracy for the BNN using the FEI database (Fig. 5(b)), which is a large effect compared to that under 1% for the BNNs using the MNIST and PubChem databases. Thus, the design of the BNN may be very important when the training database size is small. Furthermore, the type of training database affects the dependence of prediction accuracy by the inhibitory nodes. An increase in the ratio of inhibitory nodes results in a remarkable decrease in the prediction accuracy of the BNNs trained by the FEI database (Fig. 5(a)), which is different from the case of the BNNs trained by the MNIST database (size: 1k); they show a slight decrease in the prediction accuracy (Fig. 4(a)).

#### **4. Conclusions**

BNNs including three brain-like factors, disconnection, additional bypass connection, and inhibitory nodes, were studied to understand the relationship between specific local connections and training database size. For a large-size training database such as the MNIST database (size: 20k), additional bypass connections improve the prediction accuracy. On the contrary, the inhibitory nodes worsen the prediction accuracy. For the small-size training MNIST database (size: 1k), irrespective of the existence of inhibitory nodes, many BNNs showed better prediction accuracy compared to the FNN. The knockout method was used for the quantitative analysis of the improved BNN, and the findings showed that the contribution of additional bypass connections to the prediction accuracy was high. Finally, we found an improvement of up to 13% in the prediction accuracy of specific BNN structures for small-sized databases. In this study, we showed the existence of optimized BNNs depending on the size and type of training databases, therefore, we are expecting the study of how to find optimized BNN structure efficiently for a given specific database in the near future.

## References

- [1] D. M. Buss, *Evolutionary Psychology: The New Science of the Mind* (5th ed.). New York, NY, USA: Taylor & Francis, 2015.
- [2] D. Hebb, *The Organization of Behavior. A Neuropsychological Theory*. New York, NY: Wiley, 1949.
- [3] D. Miner and J. Triesch, "Plasticity-driven self-organization under topological constraints accounts for non-random features of cortical synaptic wiring," *PLoS Comput. Biol.* vol. 12, pp. e1004759, 2016.
- [4] H. Markram, "A network of tufted layer 5 pyramidal neurons," *Cerebral cortex* vol. 7, pp. 523–533, 1997.
- [5] S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, "Highly nonrandom features of synaptic connectivity in local cortical circuits," *PLoS biology* vol. 3, pp. e68, 2005.
- [6] R. Perin, T. K. Berger, and H. Markram, "A synaptic organizing principle for cortical neuronal groups," *Proc. Natl. Acad. Sci. USA* vol. 108, pp. 5419–5424, 2011.
- [7] P. L. Greer and M. E. Greenberg, "From synapse to nucleus: calcium-dependent gene transcription in the control of synapse development and function," *Neuron* vol. 59, pp. 846–860, 2008.
- [8] M. Buffelli, R. W. Burgess, G. Feng, C. G. Lobe, J. W. Lichtman, and J. R. Sanes, "Genetic evidence that relative synaptic efficacy biases the outcome of synaptic competition," *Nature* vol. 424, pp. 430–434, 2003.
- [9] R. Shapley, M. Hawken and D. L. Ringach, "Dynamics of orientation selectivity in the primary visual cortex and the importance of cortical inhibition," *Neuron*. vol. 38, pp. 689–699, 2003.
- [10] T. C. Jacob, S. J. Moss, and R. Jurd, "GABAA receptor trafficking and its role in the dynamic modulation of neuronal inhibition," *Nat. Rev. Neurosci.* Vol. 9, pp. 331–343, 2008.
- [11] D. A. Lewis, T. Hashimoto, and D. W. Volk, "Cortical inhibitory neurons and

schizophrenia,” *Nat. Rev. Neurosci.* vol. 6, pp. 312-324, 2005.

[12] Y. LeCun, Y. Bengio and G. Hinton, “Deep learning,” *Nature*. vol. 521, pp. 436–444, 2015.

[13] A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet classification with deep convolutional neural networks,” In *Proc. Adv. Neural Inf. Process. Syst.* vol. 25, pp. 1090–1098, 2012.

[14] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, “Deep neural networks for acoustic modeling in speech recognition,” *IEEE Signal Processing Mag.* vol. 29, pp. 82–97, 2012.

[15] M. Bojarski, D. D. Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, X. Zhang, J. Zhao, and K. Zieba, “End to end learning for self-driving cars,” Preprint at <https://arxiv.org/abs/1604.07316>, 2016.

[16] T. D. Albright, Cortical processing of visual motion. In F. A. Miles & J. Wallman (Eds.), *Visual motion and its role in the stabilization of gaze*, pp. 177-201, Amsterdam: Elsevier Science, 1993.

[17] L. R. Squire, D. Berg, F. E. Bloom, S. du Lac, A. Ghosh, and N. C. Spitzer, *Fundamental Neuroscience* (4<sup>th</sup> Ed.), pp. 46, Elsevier Inc., 2013.

[18] J. D. Rothstein, C. A. Pardo, L. A. Bristol, L. Jin, R. W. Kuncl, Y. Kanai, M. A. Hediger, Y. Wang, J. P. Schielke and D. F. Welty, “Knockout of glutamate transporters reveals a major role for astroglial transport in excitotoxicity and clearance of glutamate,” *Neuron*. vol. 16, pp. 675–686, 1996.

[19] J. Z. Tsien, D. F. Chen, D. Gerber, C. Tom, E. H. Mercer, D. J. Anderson, M. Mayford, E. R. Kandel and S. Tonegawa, “Subregion- and cell type–restricted gene knockout in mouse brain,” *Cell*. vol. 87, pp. 1317–1326, 1996.

[20] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, “Gradient-based learning applied to document recognition,” *Proc. IEEE*. vol. 86, pp. 2278–2324, 1998.



[21] C. E. Thomaz and G. A. Giraldi, “A new ranking method for principal components analysis and its application to face image analysis,” *Image Vis. Comput.* vol. 28, pp. 902–913, 2010.

[22] PubChem. The PubChem Project, 2015 (released in 2004). [Online]. Available: <http://pubchem.ncbi.nlm.nih.gov>

[23] M. S. Park, I. Park, Y.-S. Kang, D. Im and S.-G. Doo, “A search map for organic additives and solvents applicable in high-voltage rechargeable batteries,” *Phys. Chem. Chem. Phys.* vol. 18, pp. 26807–26815, 2016.

### **Competing financial interests**

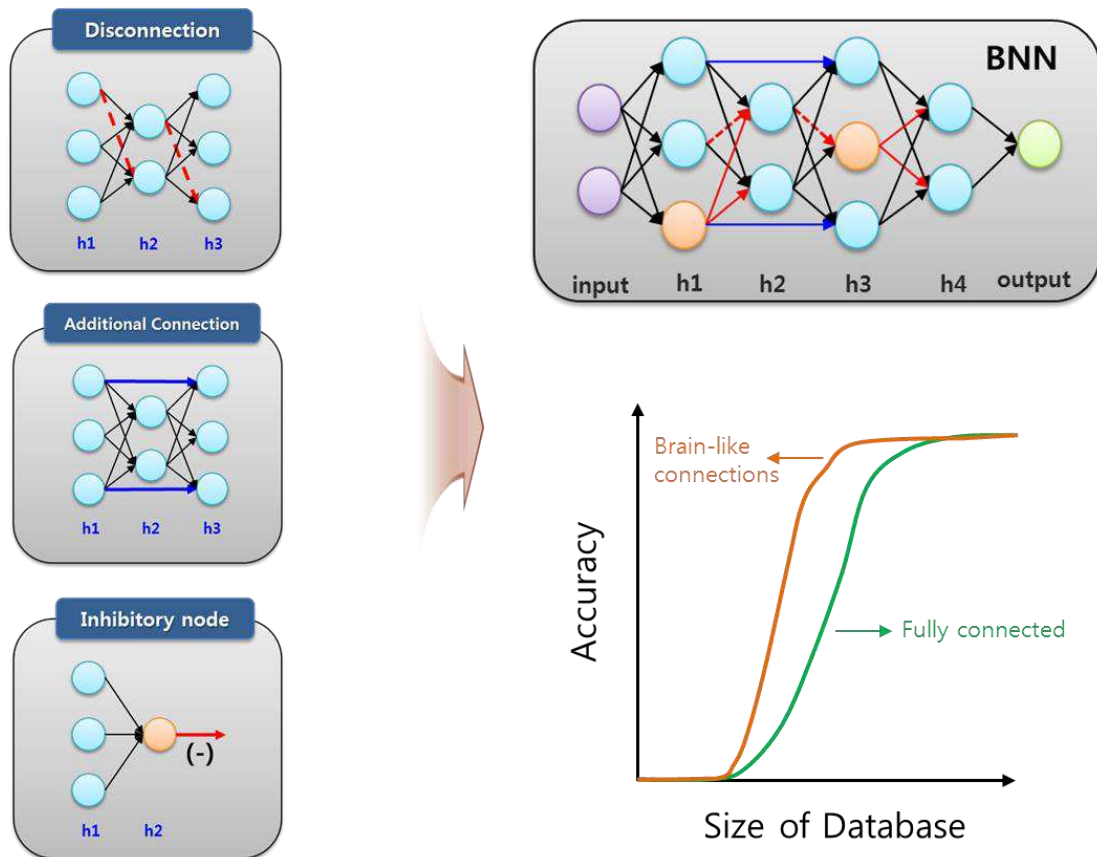
The author declares no competing financial interests.

**Table 1. Specifications of the training conditions and BNN structures**

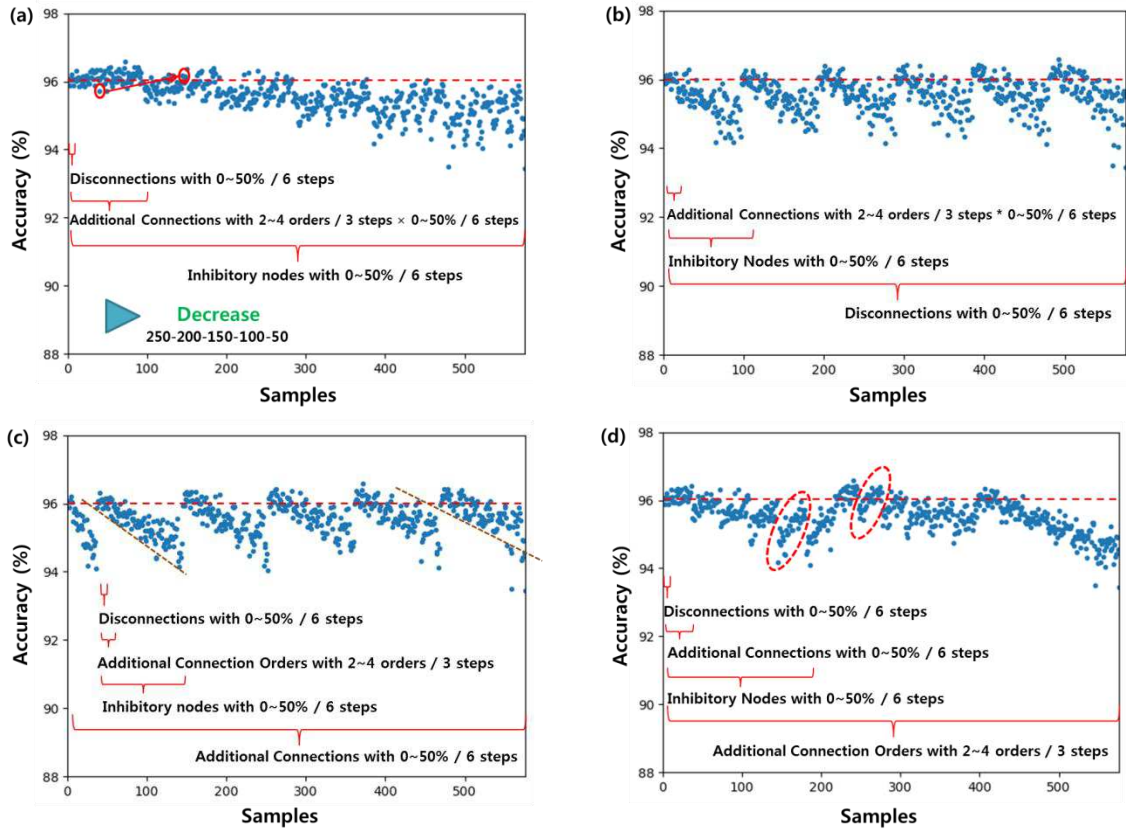
Parameters	Size	Number of cases	BNN factors	Range (interval)	Number of cases
<b>Hidden layers</b>	5	1	<b>Disconnection</b>	0–50 (10)	6
<b>Node shape</b>	Decrease	1	<b>Additional bypass connection</b>	0–50 (10), 3 levels	18
<b>Training/Test database</b>	20k/10k (MNIST) 390/30 (FEI) 20k/56,040 (PubChem)	3	<b>Inhibitory node</b>	0–50 (10)	6
<b>Epochs</b>	500	1	<b>Total cases (deduplication)</b>		<b>576</b>

**Table 2. Characteristics of the three databases used in this study**

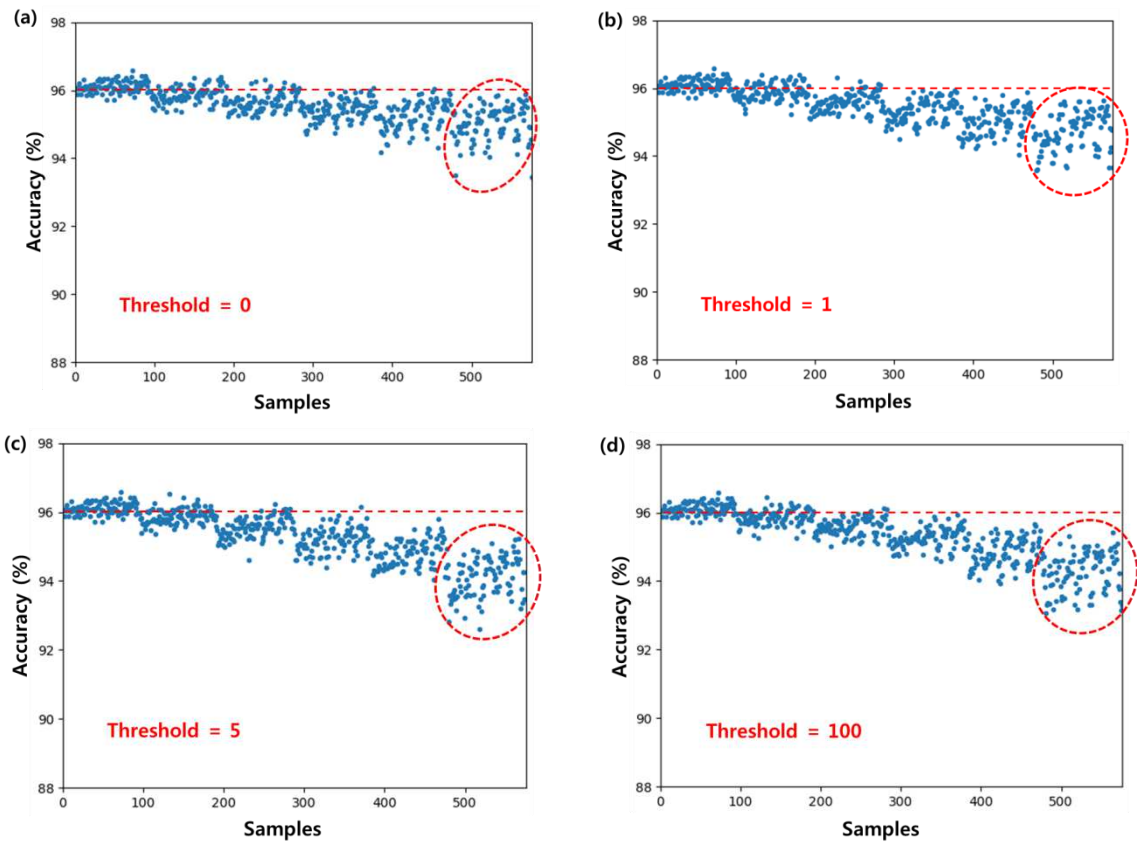
Database	MNIST	FEI	PubChem
<b>Type</b>	Handwritten Digit Image	Face Image	Chemical Compounds database
<b>Dimension</b>	28 × 28	640 × 480 × 3	1 × 86
<b>Number of Data</b>	60,000 (Training) 10,000 (Test)	2,800 (200 persons × 14 directions)	93.9 million (Compounds)



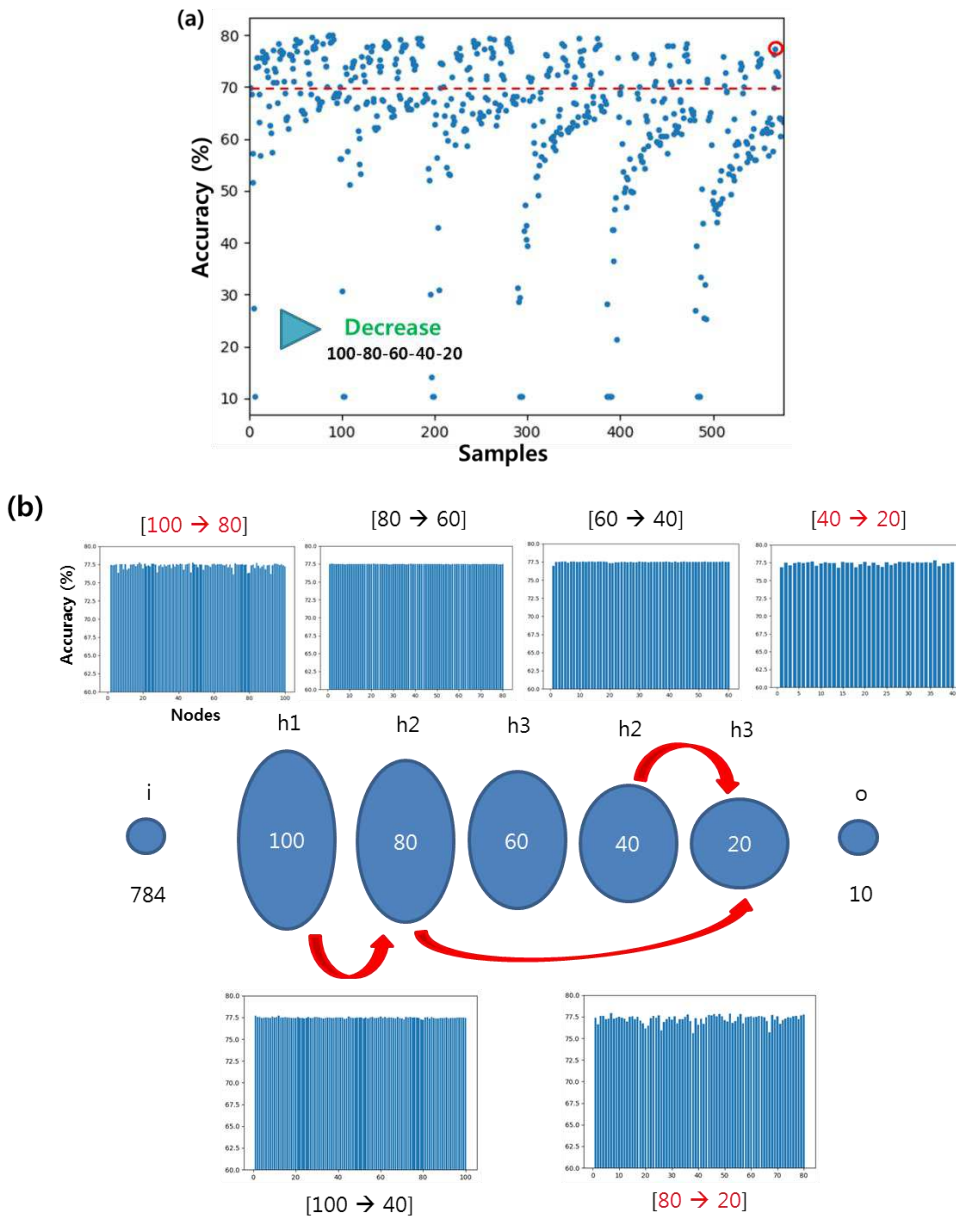
**Figure 1. (Left) Three brain-like factors included in the FNNs for the BNN test. Red dotted arrows, blue arrows, and a red arrow denote disconnected connections, additional bypass connections, and an inhibitory connection, respectively. An orange circle denotes an inhibitory node. (Right) Schematics of the BNN structure with three brain-like factors and the relationship between DB size and BNNs.**



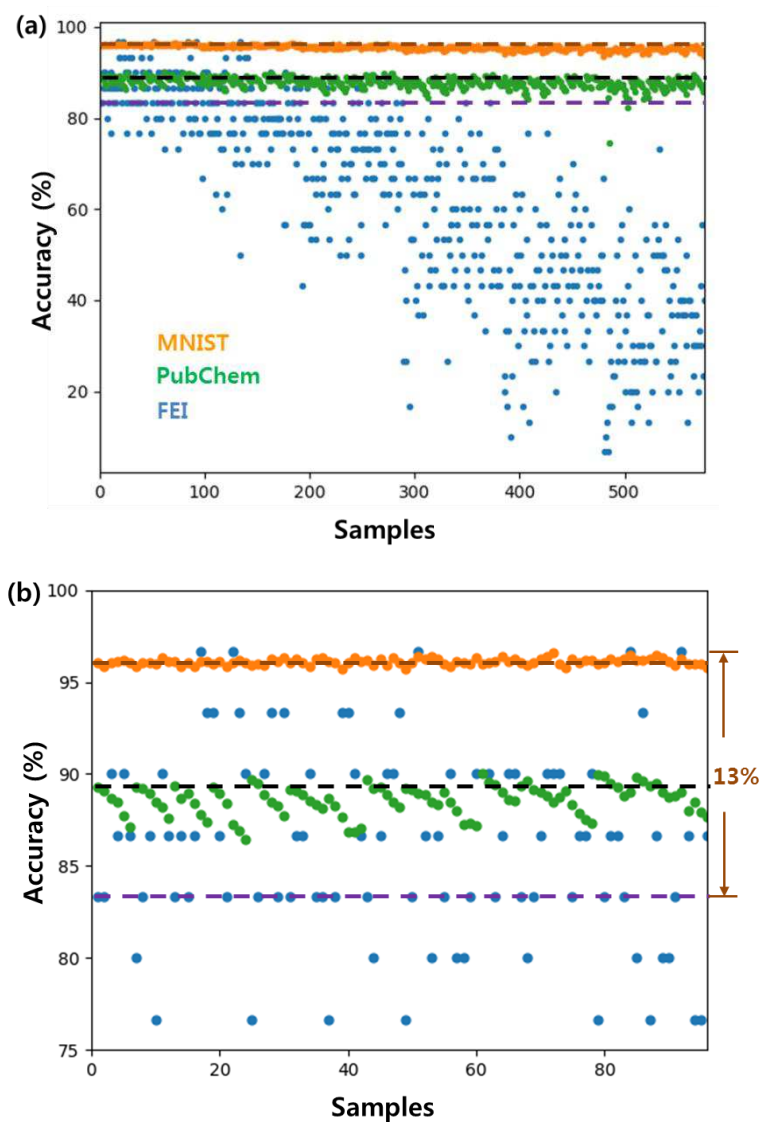
**Figure 2. Prediction accuracy for 576 BNN samples using the MNIST database. The sample order is rearranged according to (a) inhibitory nodes, (b) disconnections, (c) additional bypass connections, and (d) additional bypass connection levels. The horizontal dotted line is the prediction accuracy of the FNN. As the number of inhibitory nodes changes, the order in the prediction accuracy is not conserved in the BNNs. For instance, the 61<sup>st</sup> BNN (red circle on the left) shows minimum accuracy (95.71%) among the 96 BNNs without inhibitory nodes (the 1<sup>st</sup> cluster). However, the 169<sup>th</sup> BNN (red circle to the right), which has the same structure as the 61<sup>th</sup> BNN, shows higher accuracy (96.14%) as well as a relatively higher order among the 96 BNNs with 10% inhibitory nodes (the 2<sup>nd</sup> cluster).**



**Figure 3.** Prediction accuracy for 576 BNN samples using the MNIST database. The accuracy depends on the activation thresholds in the inhibitory nodes. The horizontal dotted line is the prediction accuracy of the FNN. The arrangement of sample order is the same as Fig. 2(a).



**Figure 4.** (a) Prediction accuracy for 576 BNN samples using the small-size MNIST database (1,000), where the horizontal dotted line is the prediction accuracy of the FNN. The arrangement of sample order is the same as Fig. 2(a). (b) The quantitative analysis of the contribution of hidden nodes to the prediction accuracy is conducted using the knockout method for the red circled sample in (a). The important connections are indicated by curved arrows.



**Figure 5. (a) Prediction accuracies for 576 BNN samples using the MNIST, FEI, and PubChem databases. (b) Prediction accuracies for 96 BNN samples without inhibitory nodes; the scale of vertical axis is enlarged from (a). The horizontal dotted line is the prediction accuracy of the FNN (red: MNIST, purple: FEI, black: PubChem). The arrangement of sample order is the same as Fig. 2(a).**



## **Supporting Information**

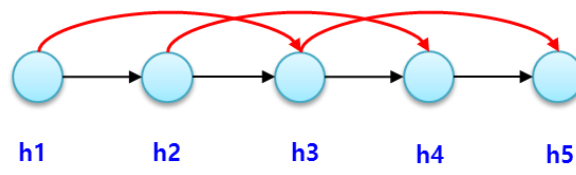
**Relationship between training database size and structural connectivity in artificial neural networks: Do specific local connections improve prediction accuracy?**

Min Sik Park\*

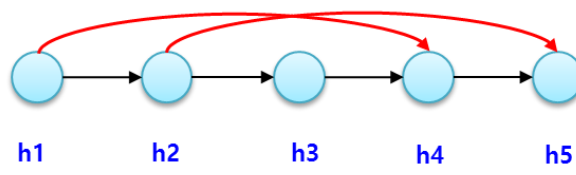
Computer-Aided Engineering Group, Samsung Advanced Institute of Technology, Samsung Electronics, 130 Samsung-ro, Suwon, Gyeonggi-do 16678, Republic of Korea

\* E-mail address: [ms91.park@samsung.com](mailto:ms91.park@samsung.com)

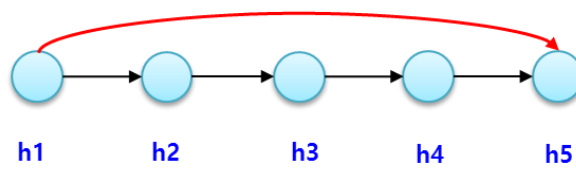
(a)



(b)



(c)



**Fig. S1 Additional bypass connection levels of additional connections, (a) level 2, (b) level 3, and (c) level 4**