IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Optimization of Neuromorphic Learning Rate Using Neural Network Simulation

Lauren Dahye Park

Korean Minjok Leadership Academy, 800 Bongwha-ro, Hoengseong-gun, Gangwon-do 25268, Republic of Korea; lauren.dahye@gmail.com

ABSTRACT: Neuromorphic computers optimized to utilize artificial intelligence are currently receiving much attention. In this study, neural network circuit simulations are performed by varying the nonlinearity and learning rate to understand better the characteristics of the individual and parallel-connected synaptic devices. The simulation results show that pattern recognition accuracy tends to decrease with increasing nonlinearity, while the accuracy is achieved more gradually when the learning rate is reduced. By comparing the simulation results with actual experimental results, I find an ideal learning rate and investigate how to optimize neuromorphic computers. These findings will help design new neuromorphic computers to determine the appropriate learning rate.

KEYWORDS: Physics and Astronomy; Theoretical; Computational and Quantum Physics; Neuromorphic Computing; Neural Network Simulation.

Introduction

The human brain can store massive amounts of information and efficiently process complex ideas. It also uses little energy and generates nearly no heat.[1] For these advantages, efforts to create special computers that think and learn like humans are advancing rapidly.[2,3] In particular, neuromorphic computers, which resemble the human brain, have recently received much attention as hardware optimized for artificial intelligence (AI).[4-9]

Current computers are composed of a central processing unit (CPU) that performs arithmetic and logic operations and a memory element that manages memory. Both of these are connected in series, as proposed by von Neumann. The separation of processing and memory increases computation speed but consumes much energy and generates heat. This paradigm raises one question: What should AI hardware look like today? This question prompted the development of neuromorphic computers.[2,3]

Furthermore, the human brain differs significantly from modern computers. In most circumstances, the human brain does not perform arithmetic operations but searches for large amounts of data or memories stored in the brain.[1] As a result, hardware capable of processing enormous amounts of big data becomes inefficient when used in a computational method designed for current tasks. People need new types of computers that look for solutions in vast data memory, such as the human brain and neuromorphic computers.

Biological synapses are necessary for transmitting nervous impulses from one neuron to another in the human brain.[10,11] Neurons specialize in transmitting signals to specific target cells, and synapses are how they accomplish this. The pre-synaptic neuron comes into close contact with the post-synaptic cell at a synapse and carries out the signaling process (Figure 1A). Synaptic weight is the strength or amplitude of a connection between two nodes.

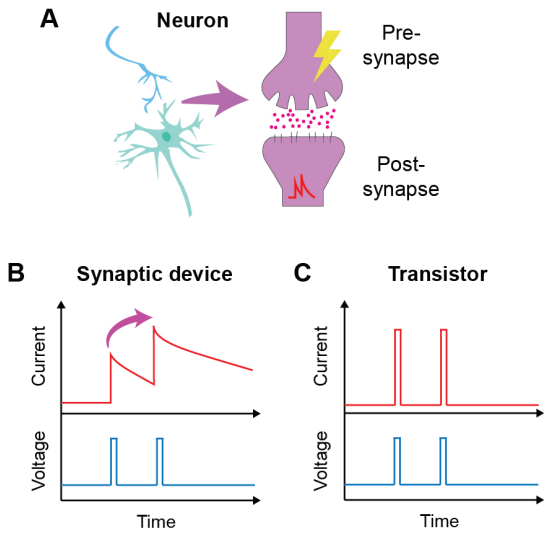

Figure 1: (A) Signal transmission in a biological neuron between pre-synapse and post-synapse; (B) Current behavior in a synaptic device in response to two voltage pulses; (C) Current behavior in a transistor in response to two voltage pulses.

On the other hand, a synaptic device, the smallest units that comprise a neuromorphic computer, is essentially a memory device for storing information. However, unlike a digital memory device that stores data in '0' and '1', it is an analog memory device that stores data in various values.[12] When two voltage pulses are applied to a synaptic device, it does not simply show two current pulses; it remembers earlier steps, so the second current pulse is added to the first and becomes larger. This weighting is a learning process (Figure 1B). The same phenomenon occurs at biological synapses. In contrast, a transistor, a representative electrical device, only computes and does not remember; therefore, it cannot perform the learning process (Figure 1C).

DOI: 10.36838/v6i8.1

In this study, I explore the characteristics of synaptic devices and perform neural network simulations based on the characteristics of a single synaptic device. In the simulations, I systematically change the 'nonlinearity', which represents specific properties of individual synaptic devices, and the 'learning rate', which determines the degree of correction for error values, and then tabulate the simulation results. Furthermore, the simulation results are applied to a real-world experimental case to determine the optimum learning rate. This will help us better understand how to optimize neuromorphic computers in the future.

Methods

When an external voltage pulse is applied continuously (multiple stimuli) in a single synaptic device, the post-synaptic current flowing through the synaptic device continues to increase (Figure 2A; red line). This memory process can be erased using a negative voltage pulse (Figure 2A; blue line). The linearity of the red and blue lines reflects the quality of a single synaptic device. The parameter 'nonlinearity' is defined as how much it deviates from the linear characteristic.

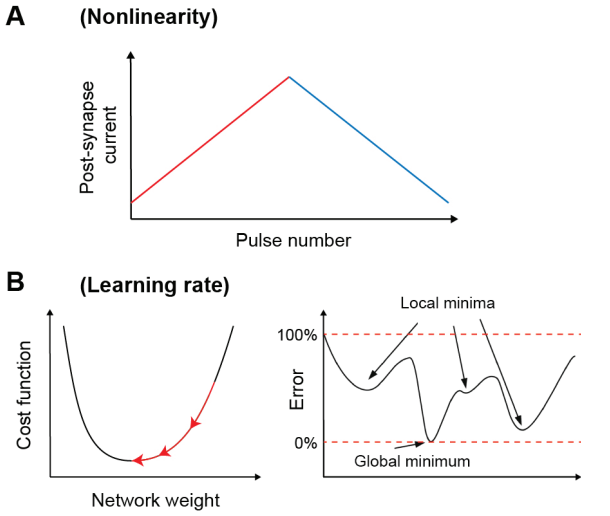

Figure 2: (A) Linear features of increasing (red line) and decreasing (blue line) post-synaptic current in a single synaptic device; (B) Schematics of the learning process of a neural network circuit based on synaptic devices connected in parallel. Procedure for determining the global minimum (left), local minima, and global minimum (right).

By connecting these synaptic devices in parallel, one can build a neuromorphic computer that learns to recognize authentic images. The 'learning rate', or the rate at which errors are subtracted from the expected output, can control the accuracy and speed at which this is learned.[2,3] The algorithm can be optimized to find the optimal global minimum by adjusting the learning rate (Figure 2B, left).

A high learning rate will quickly identify where the loss function's change rate is zero, but it is more likely to locate a local minimum, which is not the ideal answer. On the other hand, a low learning rate will look closely at the rate of change of the loss function; while it is more likely to find the point where the rate of change is zero, it is also more likely to see the global minimum, which is the genuine optimal solution (Figure 2B, right).

These two parameters, nonlinearity and learning rate, affect the accuracy, which is how well a simulation of handwritten digit recognition matches the actual digits. In the simulation, a parallel synaptic device learns approximately 3800 images of the digits 0 through 9 at 8 x 8 pixels. It delivers a probability of how accurately it can recognize the digits when presented with around 1700 new images of the digits.

To run a simulation, I used nonlinearity and learning rate as input parameters and examined how much accuracy changed. The simulations were carried out using the open-source code CrossSim.[13-16] CrossSim consists of a long and complex code, and I updated the code that determines nonlinearity and learning rate, among other parameters.

First, I ran simulations with four different nonlinearity values (0.0, 0.5, 1.0, and 2.0) to represent the properties of a single synaptic device and three different learning rates (0.1, 0.5, and 0.9). Here, nonlinearity means that the closer to 0, the more linear the single synaptic device is, and the larger the value, the more nonlinear it is. The learning rate implies that the lower the learning rate, the greater the recognition accuracy after several learning rounds, whereas the higher the learning rate, the lower the recognition accuracy.

In particular, the lower the nonlinearity of the device, the more linear the change in current weight and the more constant the rate of change. Therefore, it will reach a higher recognition rate with a shorter learning period more efficiently, even if a high learning rate is used.

In this study, systematic simulations were performed to determine the optimal learning rate by only changing the nonlinearity while keeping all other parameters constant. The number of learning processes, or epoch number, was set to a reasonably large value, 40.

Result and Discussion

1. Simulation results:

Figure 3 shows simulation results when the nonlinearity and learning rate are varied. In each of the 12 graphs, the learning rate increases as one moves to the right, and the nonlinearity rises as one moves down.

The x-axis is the epoch number, and the y-axis is the accuracy. These graphs show the accuracy achieved in a few training processes while recognizing handwritten digits in an 8 x 8 array. The epoch number denotes the number of training processes, and the accuracy denotes the recognition accuracy.

For example, in the three graphs in the first row, the nonlinearity is fixed at the smallest value of 0.0. This is based on the assumption that a single synaptic device behaves linearly. The final accuracy values as a function of learning rate are 95.725%, 96.939%, and 95.993% due to the excellent device characteristics. These results are above 95%, and it isn't easy to notice a substantial difference. However, the number of epochs required to achieve the final accuracy varies. When the learning rate is 0.1, the final accuracy is reached relatively slowly and gradually; however, when the learning rate is 0.5, the final accuracy is reached quickly at a smaller epoch number. This

trend is considerably more apparent for a learning rate of 0.9, as the final accuracy is obtained immediately at a smaller epoch number. Thus, the final accuracy is rapidly obtained at a smaller epoch when the learning rate increases.

A different trend is shown for nonlinearity values above 0.0. Suppose the learning rate is changed to 0.1, 0.5, and 0.9 when the nonlinearity is 0.5. In that case, the final accuracy decreases slightly as the learning rate increases to 94.658%, 93.656%, and 92.376%, but not significantly. The nonlinearity of 0.5 reaches the final accuracy faster than the nonlinearity of 0.0.

When nonlinearity becomes 1.0, the final accuracies are 93.100%, 90.874%, and 88.759% for learning rates of 0.1, 0.5, and 0.9, respectively, much lower than when nonlinearity is 0.5.

When the nonlinearity is increased to 2.0, the accuracy decreases as the learning rate increases. When the learning rate is increased to 0.1, 0.5, and 0.9, the final accuracy decreases significantly to 89.093%, 79.355%, and 63.606%, respectively. Nonlinearity 2.0 means that the characteristics of single synaptic devices could be better. When these devices are connected to perform neuromorphic computations, a slight change in the learning rate can cause a significant difference in image recognition accuracy. When the characteristics of a single synaptic device are poor, the learning rate must be reduced. However, this takes a long time, and the accuracy is increased by examining all local minima and obtaining the correct global minimum. In other words, to recognize accurate images quickly, a single synaptic device with strong linearity is required.

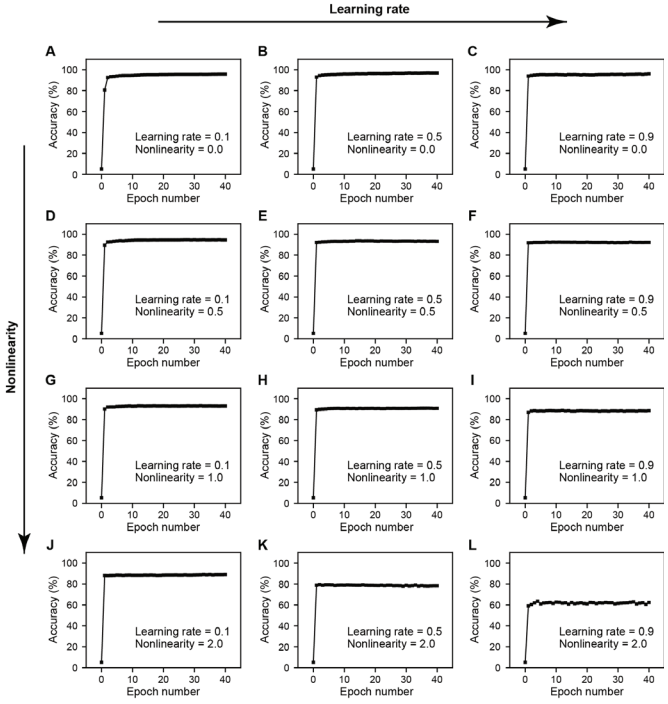

Figure 3: (A)-(L) Simulation results of accuracy as a function of epoch number, with varying nonlinearity and learning rate. In each graph, the learning rate increases to the right, whereas nonlinearity increases to the downward.

2. Analysis of simulation results:

I examine a few cases from the overall calculation results. If I zoom in on the calculation results when the learning rate is 0.5, and the nonlinearity increases, I can see that the overall accuracy decreases significantly (Figure 4). It is noted that when the performance of individual synaptic devices decreases, the pattern recognition algorithm cannot improve the accuracy beyond a specific value, even with a low learning rate.

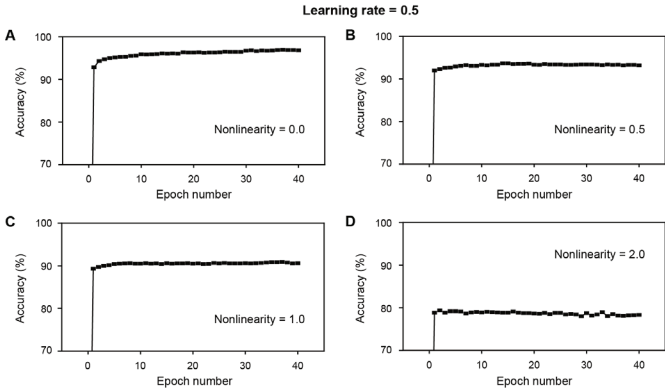

Figure 4: Simulation results of accuracy for the learning rate of 0.5, with varying nonlinearity of (A) 0.0, (B) 0.5, (C) 1.0, and (D) 2.0.

Suppose the accuracy is calculated with an increasing learning rate when the nonlinearity is 0.5. In that case, the final accuracy is not significantly different, but as the learning rate increases, the final accuracy is reached faster at a smaller epoch number (Figure 5). Once the performance of individual synaptic devices has been determined, it becomes important to decide how efficient the pattern recognition learning rate is. Because it is only sometimes necessary to learn at a low learning rate, selecting a relatively high learning rate becomes more important.

Table 1 describes all of these findings and displays the behaviors in nonlinearity and learning rate. Furthermore, once the nonlinearity of individual synaptic devices is identified, the ideal learning rate can be found quickly.

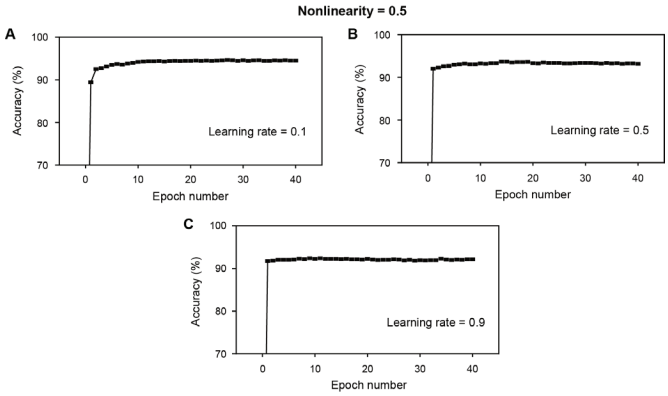

Figure 5: Simulation results of accuracy for the nonlinearity of 0.5, with varying learning rates of (A) 0.1, (B) 0.5, and (C) 0.9.

DOI: 10.36838/v6i8.1

Table 1: Simulation results of accuracy for the nonlinearity of (0.0, 0.5, 1.0, and 2.0) and learning rate of (0.1, 0.5, and 0.9).

Nonlinearity \ Learning rate	0.1	0.5	0.9
0.0	95.725 %	96.939 %	95.993 %
0.5	94.658 %	93.656 %	92.376 %
1.0	93.100 %	90.874 %	88.759 %
2.0	89.093 %	79.355 %	63.606 %

3. Comparison with experimental results:

Humans process more than 70% of their information visually. On the other hand, traditional synaptic devices use electricity as an input and must be coupled to a CMOS camera to process visual information. Light-powered synaptic devices are significantly more efficient than electricity-powered ones because they can automatically compute and store visual information, bypassing the intermediate step.[17-19]

Recently, a new photonic device that utilizes light to erase information was introduced based on a silicon nanowire with nanopores.[20] Current increases when a voltage is applied to the device to create a path for current flow, and current decreases when light blocks the path. The first process mimics the creation of memory, whereas the second mimics the erasure of memory.

In particular, this work showed the simulation results of recognizing handwritten digits using the current weighting of the light-actuated synaptic device (Figure 6). The numeric case indicates the simulation result of the virtual ideal device, which shows an accuracy of 95.5% with a learning rate of 0.1. The photonic device had a nonlinearity of 0.37, and the learning rate was simulated as 0.1. As a result, an accuracy of approximately 94.7% was achieved, nearly identical to the ideal condition.

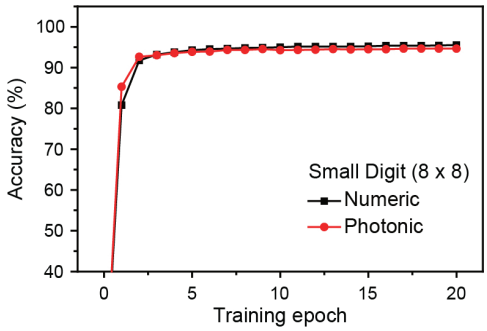

Figure 6: Simulated backpropagation training accuracies for 8×8 pixels handwritten digit images. The ideal numeric (black) and photonic processes (red) graphs are plotted as a function of the training epoch and adapted from ref. 20, copyright, 2023.

However, based on our calculations in Table 1, the nonlinearity of the photonic synaptic device, 0.37, is relatively low. Therefore, the paper of ref. 20 utilized a number too low for the learning rate, 0.1. By optimizing the learning rate, it is possible to attain comparable levels of accuracy through rapid training.

Table 1 shows that when the nonlinearity is 0.5, there is no significant difference in accuracy between learning rates of 0.1 and 0.5. Figure 5A-B, on the other hand, illustrates that the number of epochs required to achieve the final accuracy is substantially smaller.

Therefore, in this study, we simulated the accuracy when the learning rate is 0.5 and the nonlinearity is 0.37 (Figure 7). This result was compared to that shown in Ref. 20. Our simulated accuracy is 94.1%, almost similar to the value of 94.7% in previous work.[6] The maximum accuracy value can be attained at a smaller epoch number.

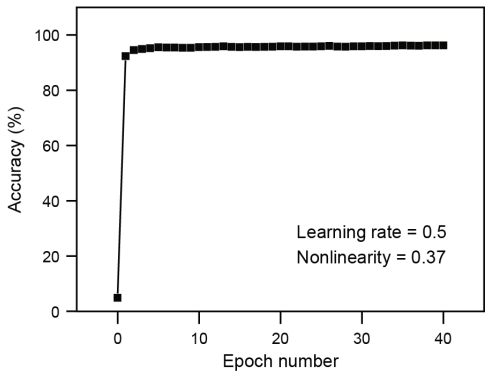

Figure 7: The simulation accuracy result for the learning rate was 0.5, and the nonlinearity was 0.37.

In the opposite case, accuracy was also simulated with a learning rate of 0.9 to examine whether it decreased significantly, as shown in Table 1 (Figure 8, red). As a result, the maximum accuracy was reached quickly, but the value was reduced to 92.9%.

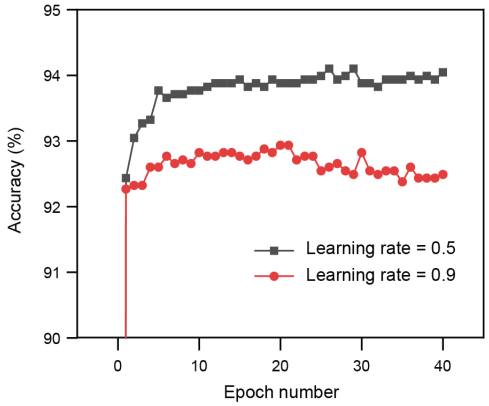

Figure 8: Simulation results of accuracy for the learning rates of 0.5 (black) and 0.9 (red). Nonlinearity is fixed at 0.37. The maximum accuracy values are 94.1% (black) and 92.9 % (red).

Overall, I found that a learning rate of 0.5 is optimal when constructing parallel circuits with light-activated synaptic devices. The results of this study will be useful in determining the ideal learning rate in a parallel neural network circuit based on a single synaptic device with a specified nonlinearity.

■ Conclusion

I performed neural network circuit simulations with different nonlinearity and learning rates to better understand the

characteristics of the individual and parallel-connected synaptic devices. The simulation results show that final accuracy decreases as nonlinearity increases, whereas final accuracy is gradually achieved when the learning rate is reduced.

A lower nonlinearity and learning rate would be desirable, but computation time would be too slow if the learning rate were too low. Since the nonlinearity of a single synaptic device is already determined in many circumstances, it is crucial to select a reasonable learning rate. To this end, I created a table summarizing image recognition accuracy with various nonlinearity and learning rates.

For example, when the learning rate is set at 0.5, and the nonlinearity increases, the accuracy decreases dramatically. When the performance of individual synaptic devices degrades with increasing nonlinearity, even with a modest learning rate, the pattern recognition algorithm cannot improve accuracy beyond a certain point. Furthermore, when the nonlinearity is adjusted to 0.5, and the learning rate is increased, the final accuracy is not much different, but it is attained faster at a lower epoch number. Once the performance of individual synaptic devices has been assessed, it is necessary to determine how efficient the learning rate is for pattern recognition.

In particular, the simulation results were utilized to evaluate the best learning rate for recently developed light-activated synaptic devices.[20] By comparing these simulation results to actual experimental data, I can determine the optimal learning rate and investigate how to optimize neuromorphic computers.

Therefore, these findings will help design new neuromorphic computers with appropriate learning rates when all other features of a single synaptic device remain fixed.

■ Acknowledgments

The author thanks Professor Kim for his valuable advice and help in my project by supporting simulation facilities.

■ References

1. Balasubramanian, V. Brain power. *Proceedings of the National Academy of Sciences* **2021**, 118, e2107022118.
2. Liu, D.; Yu, H.; Chai, Y. Low-power computing with neuromorphic engineering. *Advanced Intelligent Systems* **2021**, 3, 2000150.
3. Sun, B.; Guo, T.; Zhou, G.; Ranjan, S.; Jiao, Y.; Wei, L.; Zhou, Y. N.; Wu, Y. A. Synaptic devices based neuromorphic computing applications in artificial intelligence. *Materials Today Physics* **2021**, 18, 100393.
4. Wan, W.; Kubendran, R.; Schaefer, C.; Eryilmaz, S. B.; Zhang, W.; Wu, D.; Deiss, S.; Raina, P.; Qian, H.; Gao, B.; Joshi, S.; Wu, H.; Wong, H.-S. P.; Cauwenberghs, G. A compute-in-memory chip based on resistive random-access memory. *Nature* **2022**, 608, 504–512.
5. Merolla, P. A.; Arthur, J. V.; Alvarez-Icaza, R.; Cassidy, A. S.; Sawada, J.; Akopyan, F.; Jackson, B. L.; Imam, N.; Guo, C.; Nakamura, Y.; Brezzo, B.; Vo, I.; Esser, S. K.; Appuswamy, R.; Taba, B.; Amir, A.; Flickner, M. D.; Risk, W. P.; Manohar, R.; Modha, D. S. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* **2014**, 345, 668–673.
6. Zidan, M. A.; Strachan, J. P.; Lu, W. D. The future of electronics based on memristive systems. *Nature Electronics* **2018**, 1, 22–29.
7. Lee, T. H.; Loke, D.; Huang, K.-J.; Wang, W.-J.; Elliott, S. R. Tailoring transient-amorphous states: towards fast and power-efficient phase-change memory and neuromorphic computing. *Advanced Materials* **2014**, 26, 7493–7498.
8. Han, M.-K. Neuromorphic chip (Y-KAST report), *The Korean Academy of Science and Technology* **2020**, 1, 1–28.
9. Han, M.-K. Machine learning (Y-KAST report), *The Korean Academy of Science and Technology* **2019**, 1, 1–32.
10. Perea, G.; Navarrete, M.; Araque, A. Tripartite synapses: astrocytes process and control synaptic information. *Trends in Neurosciences* **2009**, 32, 421–431.
11. Lee, B.-U. Biology glossary for high school students, ShinWon Books, 2002.
12. Mandal, S.; El-Amin, A.; Alexander, K.; Rajendran, B.; Jha, R. Novel synaptic memory device for neuromorphic computing. *Scientific Reports* **2014**, 4, 5333.
13. van de Burgt, Y.; Lubberman, E.; Fuller, E. J.; Keene, S. T.; Faria, G. C.; Agarwal, S.; Marinella, M. J.; Talin, A. A.; Salleo, A. A non-volatile organic electrochemical device as a low-voltage artificial synapse for neuromorphic computing. *Nature Materials* **2017**, 16, 414–418.
14. Agarwal, S.; Plimpton, S. J.; Hughart, D. R.; Hsia, A. H.; Richter, I.; Cox, J. A.; James, C. D.; Marinella, M. J. Resistive memory device requirements for a neural algorithm accelerator. *2016 International Joint Conference on Neural Networks (IJCNN)* **2016**, 929–938.
15. Spear, M.; Kim, J. E.; Bennett, C. H.; Agarwal, S.; Marinella, M. J.; Xiao, T. P. The impact of analog-to-digital converter architecture and variability on analog neural network accuracy. *IEEE Journal on Exploratory Solid-State Computational Devices and Circuits*, **2023**, 9, 176–184.
16. Xiao, T. P.; Feinberg, B.; Bennett, C. H.; Prabhakar, V.; Saxena, P.; Agrawal, V.; Agarwal, S.; Marinella, M. J. On the accuracy of analog neural network inference accelerators. *IEEE Circuits and Systems Magazine* **2022**, 22, 26–48.
17. Wang, G.; Wang, R.; Kong, W.; Zhang, J. Simulation of retinal ganglion cell response using fast independent component analysis. *Cognitive Neurodynamics* **2018**, 12, 615–624.
18. Zhou, F.; Zhou, Z.; Chen, J.; Choy, T. H.; Wang, J.; Zhang, N.; Lin, Z.; Yu, S.; Kang, J.; Wong, H.-S. P.; Chai, Y. Optoelectronic resistive random access memory for neuromorphic vision sensors. *Nature Nanotechnology* **2019**, 14, 776–782.
19. Lee, K.-J.; Kim, J. H.; Jeon, S.; Shin, C. W.; Kim, H.-R.; Park, H.-G.; Kim, J. Polarization-dependent memory and erasure in quantum dots/graphene synaptic devices. *Nano Letters* **2024**, 24, 2421–2427.
20. Lee, H.-C.; Kim, J.; Kim, H.-R.; Kim, K.-H.; Park, K.-J.; So, J.-P.; Lee, J. M.; Hwang, M.-S.; Park, H.-G. Nanograin network memory with reconfigurable percolation paths for synaptic interactions. *Light: Science & Applications* **2023**, 12, 118.

■ Author

Lauren Dahye Park attends the Korean Minjok Leadership Academy (G10) in South Korea.

DOI: 10.36838/v6i8.1