Applied Physics Express

LETTER • **OPEN ACCESS**

Random network device fabricated using Ag_2Se nanowires for data augmentation with binarized convolutional neural network

OPEN ACCESS

Applied Physics Express 16, 014002 (2023) LETTER

https://doi.org/10.35848/1882-0786/acae6a

Random network device fabricated using Ag_2Se nanowires for data augmentation with binarized convolutional neural network

Takumi Kotooka[1], Yuichiro Tanaka[2], Hakaru Tamukoh[1,2], Yuki Usami[1,2*], and Hirofumi Tanaka[1,2*]

[1]Department of Human Intelligence System, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology (Kyutech), 2-4 Hibikino, Wakamatsu, Kitakyushu 808-0196, Japan

[2]Research Center for Neuromorphic AI Hardware, Kyushu Institute of Technology (Kyutech) 2-4 Hibikino, Wakamatsu, Kitakyushu 808-0196, Japan

*E-mail: usami@brain.kyutech.ac.jp; tanaka@brain.kyutech.ac.jp

Received November 14, 2022; revised December 8, 2022; accepted December 25, 2022; published online January 12, 2023

An Ag_2Se nanowire random network was fabricated for application as a data augmentation device and combined with a binary convolutional neural network (BCNN) to achieve high accuracy in voice classification tasks. Due to the nonlinear high-dimensional characteristics resulting from the formation of the conductive filament at the cross junction, the Ag_2Se device could transform input data into higher-order multiple signals, thereby enhancing the accuracy of the classification task by augmenting input signals. The results indicate that materials can realize data augmentation with the same performance as software, suggesting that material-based hardware can be used as an elemental technology for information processing. © 2023 The Author(s). Published on behalf of The Japan Society of Applied Physics by IOP Publishing Ltd

Research on artificial intelligence (AI) technologies, such as machine learning and artificial neural networks (ANNs), has progressed rapidly in recent years. Specialized models for each class of tasks, such as convolutional and recursive, have been developed, and these models have been proven to have high accuracy.[1–6] However, the results of such tasks can be obtained only when a sufficient amount of data is provided. Regardless of the quality and accuracy of the model, it will not perform satisfactorily if the data are insufficient. The data augmentation method is used to address this problem. Data augmentation is used in data preprocessing to virtually increase the volume of data from a small amount of data that is originally available. Some of the techniques used to accomplish this include adding noise to the original data or using data transformation operations, such as shifting and stretching.[7,8] Although this technique can compensate for missing data, the preprocessing can incur high costs and requires an efficient system. As the volume of augmented data increases, similar data obtained by adding noise or other alternative signals to the original image can be obtained from a variety of data by filtering the signals. Material-based devices have been used to implement this functionality in hardware components.[9–15] These devices use the physical properties of materials, including nonlinear and high-dimensional properties, to process information. We have performed benchmark tasks, such as voice classification and image recognition to demonstrate their information processing capabilities,[16–18] and they are expected to be applied to data augmentation filters to perform various transformations of the input. Some material-based information processing devices that use atomic switch networks have been reported, and their nonlinear and high-dimensional properties have been theoretically and experimentally confirmed.[19–21] Because Ag_2Se also exhibits atomic switching phenomena driven by the redox transformations of silver atoms, we propose the use of Ag_2Se nanowire random network devices in combination with atomic switch networks as data extension devices.[22–24] The material also has high electrical conductivity, low lattice thermal conductivity, and high thermal stability. In our previous study, we also found that Ag_2Se nanowire random networks have electrically nonlinear due to atomic switch elements formed at the junctions of the nanowires and high-dimensional properties due to the network geometry,[25] which could be applied to the nonlinear conversion of signals and augmentation of the quantity of data.

After augmentation, the data need to be inputted into an ANN for training. Material-based data augmentation devices need to be integrated with the neural network circuit in a system configuration. To decrease the power requirement for computation, a binary convolutional neural network (BCNN) was employed to confirm compatibility with hardware-oriented neural networks. A BCNN is a network in which weights between the convolution layer and the max pooling layer and affine layer are represented as binary values.[26–29] In addition, because the BCNN uses binary values, data processing can be realized using digital logic circuits, which allows for easy hardware implementation compared to other neural network models. Voice classification based on CNN is highly accurate,[30] and BCNN has already been implemented in hardware using field programmable gate arrays.[31–33] Hence, BCNNs were used in the present study. The objective was to use an Ag_2Se nanowire random network device as a data augmentation device in conjunction with a BCNN to construct a highly efficient hardware-friendly computing system. The Ag_2Se nanowires were synthesized and analyzed using X-ray diffraction (XRD) and scanning electron microscopy (SEM). The electrical properties of the Ag_2Se nanowire random network device were measured and various output signals were obtained from the random networks. Then, voice classification was performed to evaluate the data augmentation performance of the device.

Ag_2Se was synthesized based on the procedure detailed in the literature.[22] Sodium selenite (0.25 g, Sigma Aldrich 99.9%) and glucose (1.5 g, Wako, 98%) were added to 100 ml deionized (DI) water at 90 °C without stirring. A brick-red precipitate was formed at the bottom after 20 min. This material was washed several times with DI water.

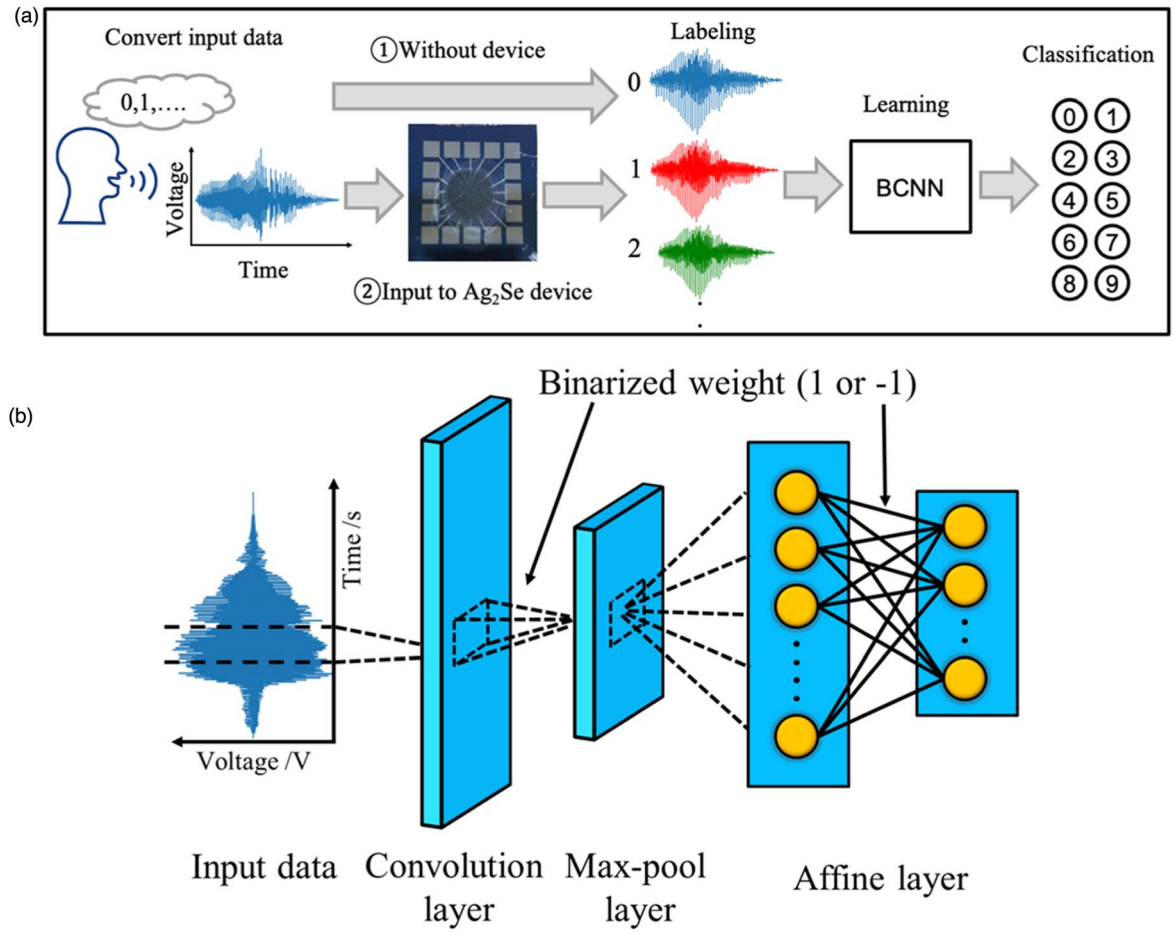

Fig. 1. (Color online) (a) Schematic diagram of voice classification procedure with Ag_2Se nanowire random network device. A two-pattern process was conducted with and without the Ag_2Se device to compare the performance. (b) Schematic BCNN components. The input was supplied to the convolutional layer and then max-pooled to rearrange for the insertion into the affine layer.

Subsequently, it was dispersed again in isopropanol and stored in a dark room at 24 °C for one week, resulting in the formation of t-Se nanowires. These nanowires were submerged in an $AgNO_3$ solution for 3 h to obtain the target material.

The random network device was fabricated using electron beam (EB) lithography. First, a Si substrate coated with SiO_2 (thickness: 200 nm) was used. The substrate was coated by resist (gL 2000: anisole = 1:1, Gluon Lab), and the spinning speed and time were 5000 rpm and 1 min. This exposed the EB (current: 30 nA), and the development of the substrate using a ZED-N50 (Zeon) solution was carried out for 5 min. The developed substrate was then sputtered with Al (thickness, 100 nm), and the sputtered substrate was lifted off using dimethyl sulfoxide solution at 60 °C for 20 min. Finally, the device was fabricated by dropping the synthesized nanowires at the electrode centers. The electronic measurements were controlled using the LabVIEW 2020 system. The input was a sinusoidal waveform (amplitude: 2 V_{pp}, frequency: 11 Hz), and 15 *V*–*t* outputs were obtained with the DAQ system (PXIe-6363). The outputs from the *V*–*t* measurements were converted by fast Fourier transformation with triangular windows to check for the generation of higher harmonics. A free-spoke-digit-data dataset[34] that included voice-recorded data for the pronunciation of the numbers zero to nine by six speakers was used to demonstrate voice classification. Each number was pronounced 50 times by each of the speakers. LabVIEW was used to convert these data to analog time-voltage signals as the input for the fabricated device. Fifteen of these outputs were measured at a sampling rate of 1000/s. After recording, the data were labeled as training signals and inputted into the BCNN model for the classification task. The processes were conducted with and without the device to verify the performance results. The procedure for the voice classification task is shown in Fig. 1. A BCNN has two layers, namely, a convolutional layer (input channel: 1, output channel: 30, filter size: 50) and an affine layer. The labeled signals were inputted to the convolutional layer, subjected to max pooling (filter size: 4), and sent to the affine layer. A stochastic gradient descent method was used for training. A total of 6750 training data points were used for number classification, which was obtained by data augmentation from 15 outputs (without case: 450). In the speaker classification, the number of training data points was 40500 (without case: 2700). The number of test data points was set to 50 (number classification) and 300 (speaker classification), respectively. The number of epochs was set to 50. The classification performance was evaluated using a confusion matrix and four indicators. The confusion matrix is composed of the actual and predicted labels from the learning model. A higher diagonal element indicates that the classification performance

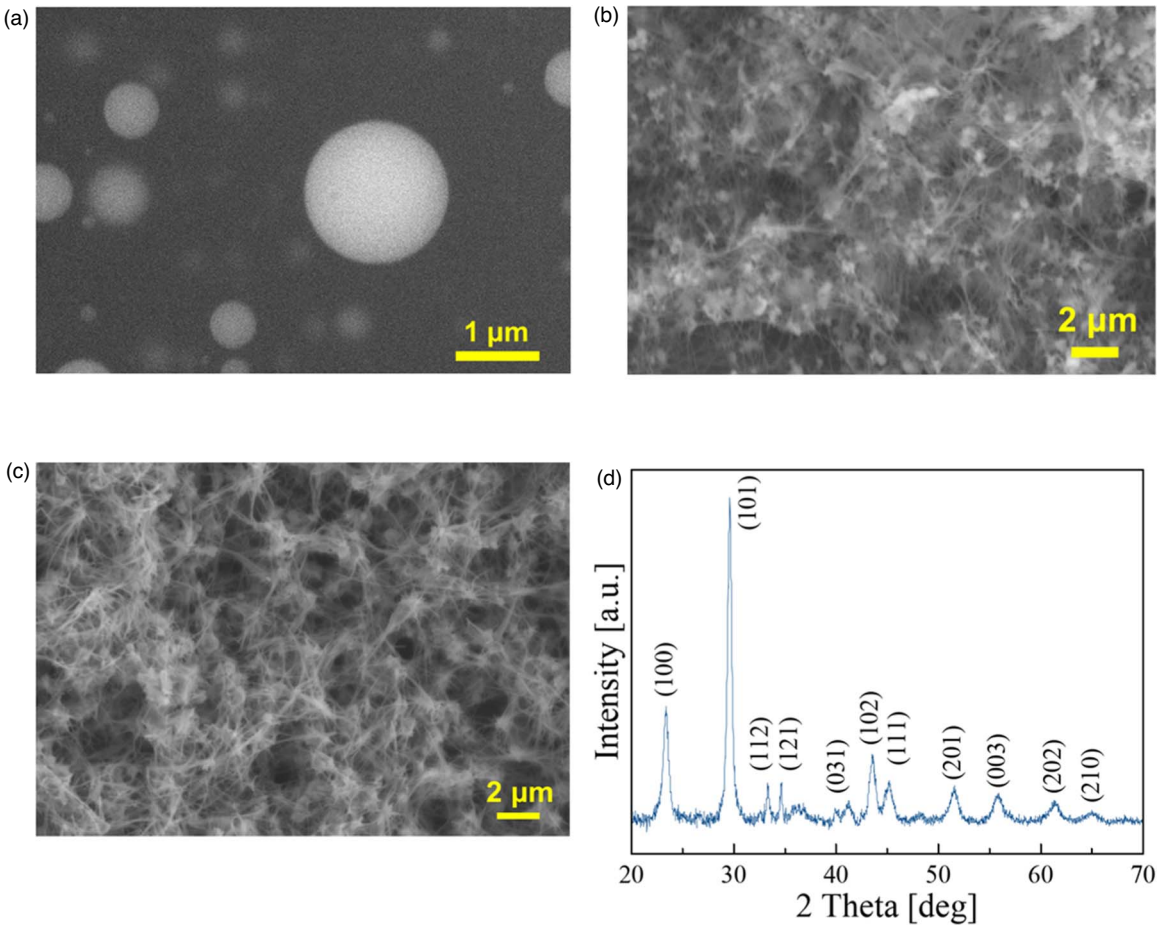

Fig. 2. (Color online) (a) SEM image of amorphous Se obtained from precipitation. (b) SEM image of synthesized Se nanowires. Amorphous Se remains in the network. (c) SEM image of Ag_2Se nanowires after submersion in $AgNO_3$ aq. The morphology does not change by the submersion. (d) XRD results of synthesized Ag_2Se nanowires.

is better. From the value of the confusion matrix, four values can be calculated as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}; \quad \text{(I)}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}; \quad \text{(II)}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}; \quad \text{(III)}$$

$$\text{F1 score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}, \quad \text{(IV)}$$

where TP, TF, NP, and NF denote the true positive, true negative, negative positive, and negative false, respectively. These values were high, indicating the success of the classification task.

Figure 2(a) shows an SEM image of the brick-red precipitate obtained in the synthesis. The results of an earlier study confirm the spherical material to be amorphous Se.[28] The nanowire was obtained one week after the precipitate was formed, as shown in Fig. 2(b). Some small amorphous Se particles were observed as the nanowires grew from it and formed t-Se nanowires.[35,36] Figures 2(b) and 2(c) show SEM images of the synthesized materials before and after immersion in the $AgNO_3$ aqueous solution, respectively. Both images clearly show the nanowire geometry, and no morphological changes were observed before and after immersion. The diameter of the nanowires was approximately 73.8 nm. Figure 2(d) shows the XRD results of the synthesized material, indicating the crystallinity of the material. The designed and fabricated Ag_2Se nanowire random network device is shown in Figs. 3(a) and 3(c). The electrodes were composed of aluminum, and the distance between the devices was 200 μm, as shown in Fig. 3(b). After casting, the center of the device was examined by SEM and confirmed to have a disordered network. The results of the *V*–*t* measurement are shown in Fig. 3(d). All the outputs have a phase shift, which indicates that all the paths from the input to the output electrodes are complicated and indicate the possibility of obtaining various inputs. Phase shifting is useful as a data augmentation technique for voice data since voices with pseudo-timing shifts can be obtained. Figure 3(e) shows a Lissajous plot for each output. The shape shows the relationship between the input and output and reveals the network dynamics. The random network device with Ag_2Se nanowire exhibited atomic switching behavior owing to the redox reaction that occurred in the network. The difference in output for each electrode could result from the different conduction pathways. By randomly forming an atomic switch, it was possible to create different outputs from a single signal; these outputs were not identical, and the variety

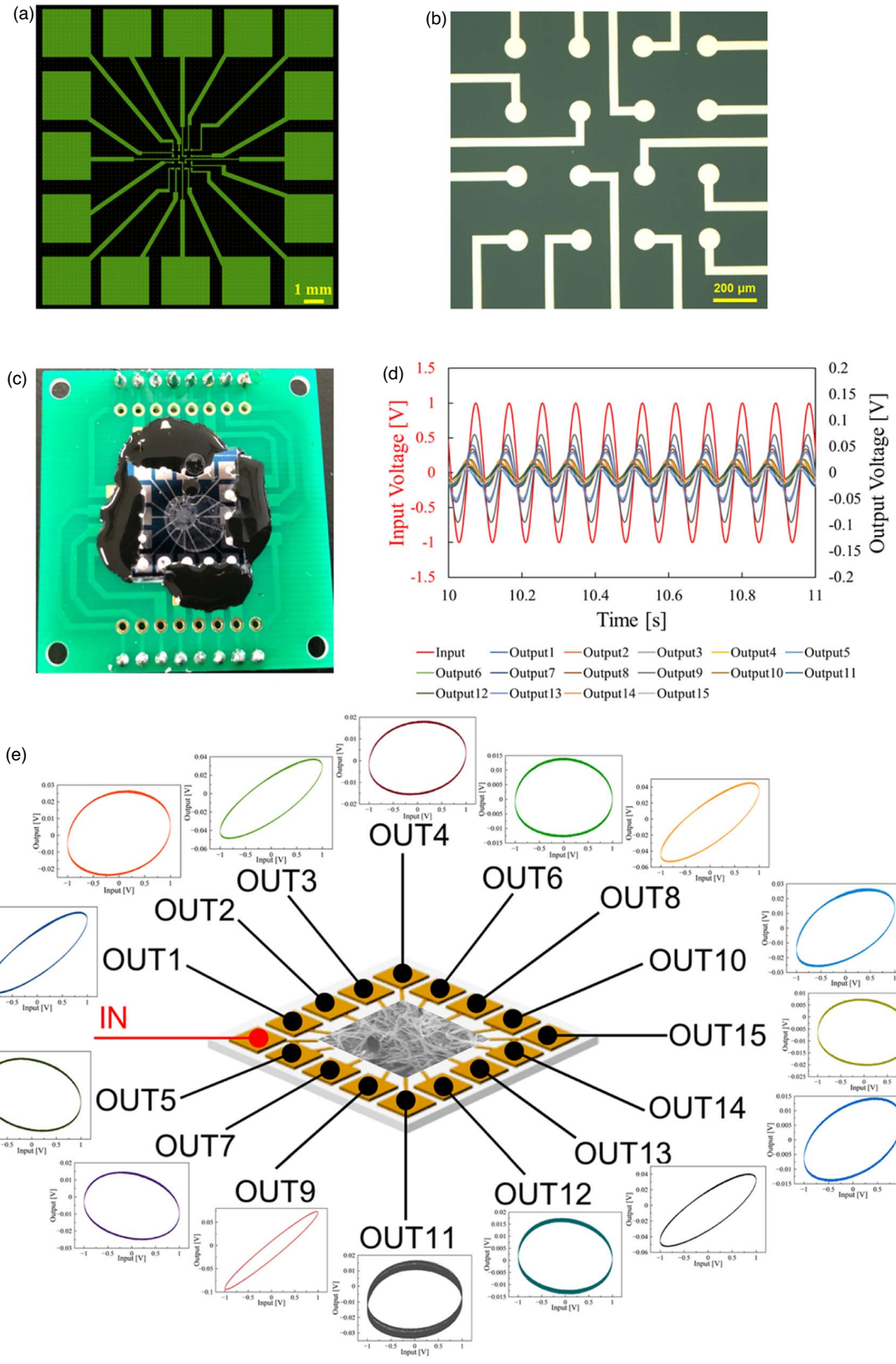

Fig. 3. (Color online) (a) Designed electrodes. The number of electrodes was 16, and the distance between electrodes at the center was 200 μm, as shown in (b). (c) Fabricated electrode chip. Center part was cast using Ag_2Se nanowire solution and formed a random network. (d) *V*–*t* results of Ag_2Se nanowire random network device. Red signal indicates the input (amplitude; 2 V_{pp}, frequency: 11 Hz). (e) Lissajous plot at each output. All plots exhibit different shapes due to different conducting paths.

available in the output was amenable to the data augmentation process.

Figures 4(a) and 4(b) show the confusion matrix of the number classification results with and without the Ag_2Se nanowire random network device. Better classification results were obtained with the Ag_2Se random network device (accuracy: 92%) than without it (accuracy: 74%). Figures 4(c) and 4(d) show the speaker classification results

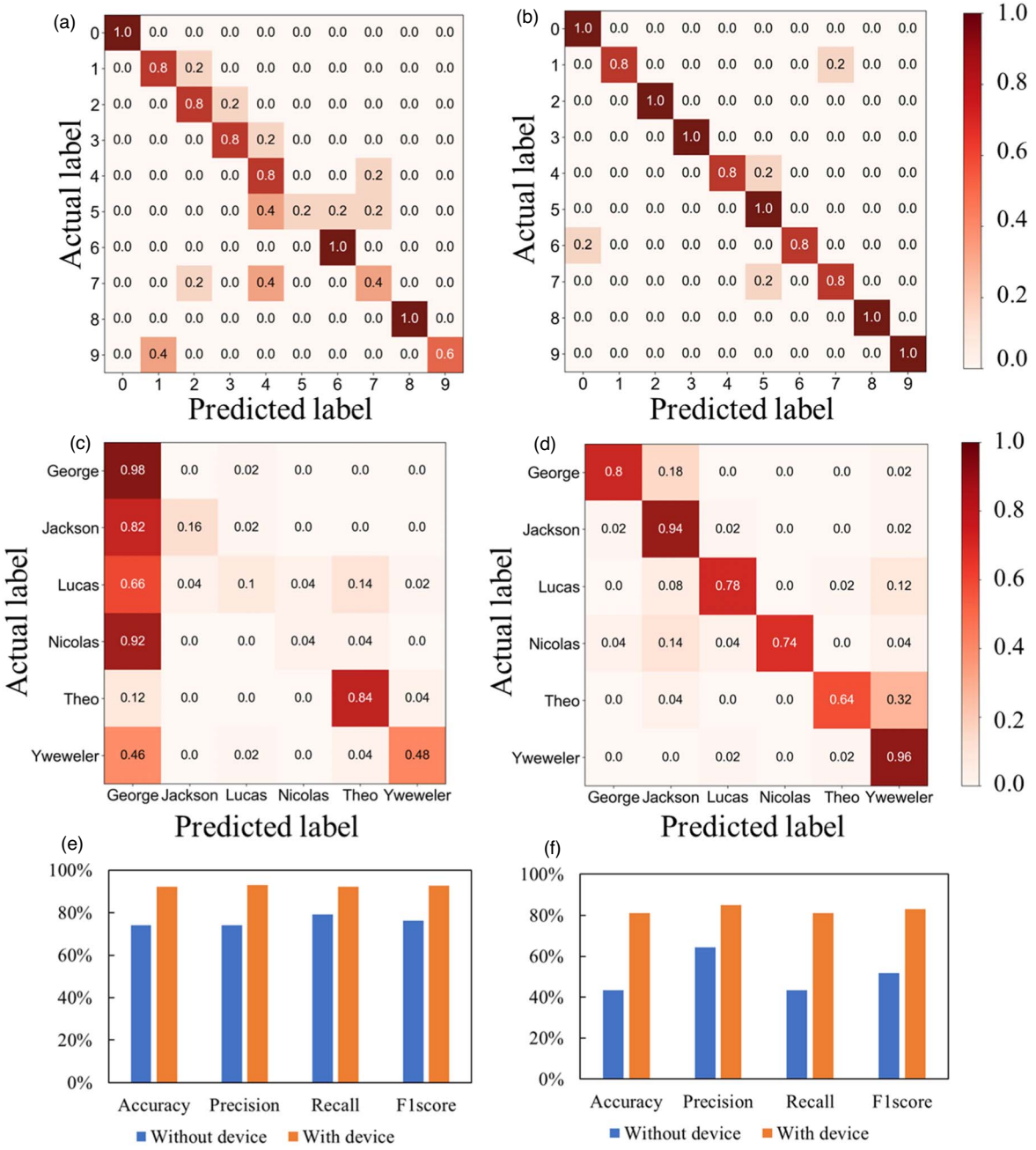

Fig. 4. (Color online) Voice classification results through the BCNN. Confusion matrices of number classification (a) without and (b) with the device. Confusion matrices of speaker classification (c) without and (d) with the device. Evaluation values comparing without and with the device; (e) number and (f) speaker classification task.

without and with the device, respectively. The accuracy obtained with the Ag_2Se nanowire random network device was 43.3% and 81% without it. These results also indicated that the Ag_2Se random network device improved the classification performance effectively. Speaker classification often requires more data than readily available data, and the results of this study indicate significant differences in accuracy between cases with and without augmentation devices. The presence of the device helped to successfully increase the volume of training data while maintaining the characteristics of each user. The four indicators are higher when the device is used compared to when not used, as shown in Figs. 4(e) and 4(f). These results have been attributed to two reasons: one is the increased availability of data, and the other is the conversion of input signals to various other types of signals owing to the nonlinear and high-dimensional properties in the Ag_2Se nanowire network. Figures 5(a) and 5(b) illustrate the relationship between the quantity of training data and classification accuracy for number classification and speaker classification, respectively. The accuracy increased as the available quantity of training data increased. Generally, as the quantity of training data for a task increases, the task results improve.[37] However, if the variety of signals in the training data is limited, this performance improvement cannot be realized. Therefore, to add characteristics such as noise and randomness, the fabricated device was developed as a filter that could realize nonlinear conversions. As a result, the variety of data used for learning increased, and high performance was achieved. The results show that material-

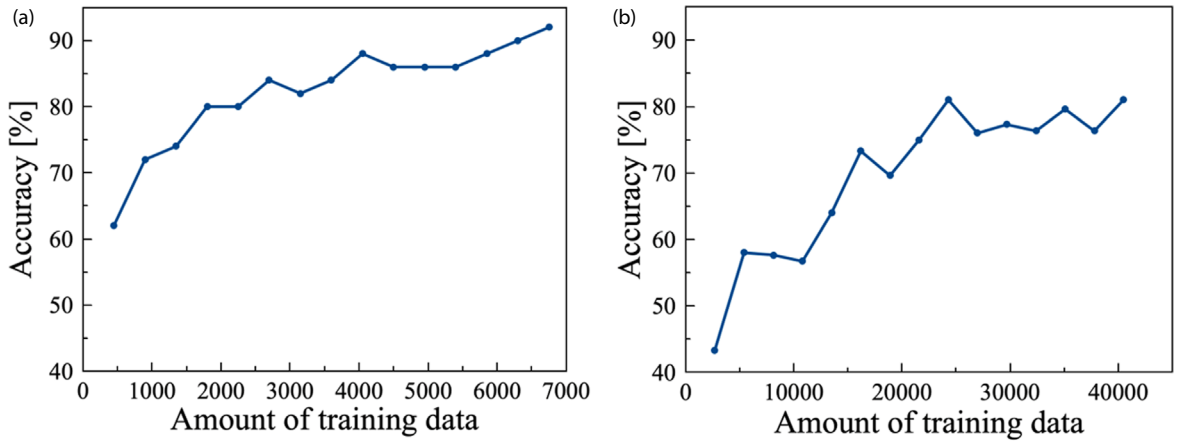

Fig. 5. (Color online) Amount of training data and accuracy dependence. (a) Number classification. (b) Speaker classification.

based device successfully converted one signal to multiple signals. The accuracy can be improved further by increasing the number of electrodes and enhancing the nonlinear and high-dimensional properties of the material. In the Ag_2Se nanowire network used in this study, nonlinear characteristics were caused by the filament formation of the atomic switch element. In addition, to obtain a variety of outputs, it is necessary to increase the number of conduction paths through many atomic switch elements. In other words, the high dimensionality may depend on the network topology of the nanowires. In this study, material-based random network devices were incorporated into preprocessing as data augmentation elements, and their usefulness was demonstrated through voice classification. In the future, such augmentation devices can possibly be incorporated into electronic circuits and combined with other AI hardware to develop highly efficient AI learning systems.

In conclusion, we demonstrated that an Ag_2Se nanowire random network device can be used as a data augmentation device. An Ag_2Se nanowire random network device was fabricated, and the electric measurements indicated that it had nonlinear properties resulting from the atomic switch network. Lissajous plots showing a variety of outputs indicated that an input signal was converted to many nonlinear and high-dimensional signals resulting in data augmentation. A demonstration of the voice classification function showed a marked performance improvement with the Ag_2Se nanowire random network device to increase the variety of data because of its nonlinear and high-dimensional properties. In the future, this device will be used in electronic circuit chips of AI systems, which is expected to facilitate a more efficient architecture.

Acknowledgments This work was financially supported by KAKENHI (Grant Nos. 19H02559, 19K22114, 20K21819, and 21K14527) and JST CREST (Grant No. JPMJCR21B5), and JST ACT-X (Grant No. JPMJAX22K4). The work was conducted at the Kitakyushu Semiconductor Center, supported by the Nanotechnology Platform Program (Nanofabrication) of the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan. Y. U. thanks Asahi Kohsan Co., Ltd. for their financial support through the Kitakyushu Foundation for the Advancement of Industry, Science, and Technology, Japan.

ORCID iDs Takumi Kotooka https://orcid.org/0000-0002-4513-1878 Yuichiro Tanaka https://orcid.org/0000-0001-6974-070X Hakaru Tamukoh https://orcid.org/0000-0002-3669-1371 Yuki Usami https://orcid.org/0000-0002-8583-325X Hirofumi Tanaka https://orcid.org/0000-0002-4378-5747

1) A. Krizhevsky, I. Sutskever, and G. E. Hinton, Commun. ACM **60**, 84 (2017).
2) H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, Procedia Comput. Sci. **90**, 200 (2016).
3) S. Ren, K. He, R. Girshick, and J. Sun, IEEE Trans. Pattern Anal. Mach. Intell. **39**, 1137 (2017).
4) A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber, ACM Int. Conf. Proceeding Ser. **148**, 369 (2006).
5) K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, BMVC 2014 - Proc. Br. Mach. Vis. Conf. 2014, Computer Vision and Pattern Recognition (cs.CV), 10.48550/arxiv.1405.3531.
6) A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, Nature **542**, 115 (2017).
7) J. Salamon and J. P. Bello, IEEE Signal Process Lett. **24**, 279 (2017).
8) D. M. Allen, Technometrics **16**, 125 (1974).
9) G. Milano, G. Pedretti, K. Montano, S. Ricci, S. Hashemkhani, L. Boarino, D. Ielmini, and C. Ricciardi, Nat. Mater. 2021 **21**, 195 (2021).
10) Y. Zhong, J. Tang, X. Li, B. Gao, H. Qian, and H. Wu, Nat. Commun. **12**, 408 (2021).
11) A. Z. Stieg, A. V. Avizienis, H. O. Sillin, C. Martin-Olmos, M. Aono, and J. K. Gimzewski, Adv. Mater. **24**, 286 (2012).
12) M. Nakajima, K. Minegishi, Y. Shimizu, Y. Usami, H. Tanaka, and T. Hasegawa, Nanoscale **14**, 7634 (2022).
13) H. Tanaka et al., Neuromorphic Comput. Eng. **2**, 022002 (2022).
14) D. Banerjee, S. Azhari, Y. Usami, and H. Tanaka, Appl. Phys. Express **14**, 105003 (2021).
15) Hadiyawarman, Y. Usami, T. Kotooka, S. Azhari, M. Eguchi, and H. Tanaka, Jpn. J. Appl. Phys. **60**, SCCF02 (2021).
16) H. Tanaka, M. Akai-Kasaya, A. TermehYousefi, L. Hong, L. Fu, H. Tamukoh, D. Tanaka, T. Asai, and T. Ogawa, Nat. Commun. 2018 **9**, 1 (2018).
17) D. Banerjee, T. Kotooka, S. Azhari, Y. Usami, T. Ogawa, J. K. Gimzewski, H. Tamukoh, and H. Tanaka, Adv. Intell. Syst. **4**, 2100145 (2022).
18) Y. Usami et al., Adv. Mater. **33**, 2102688 (2021).
19) K. S. Scharnhorst, J. P. Carbajal, R. C. Aguilera, E. J. Sandouk, M. Aono, A. Z. Stieg, and J. K. Gimzewski, Jpn. J. Appl. Phys. **57**, 1 (2018).
20) E. C. Demis, R. Aguilera, H. O. Sillin, K. Scharnhorst, E. J. Sandouk, M. Aono, A. Z. Stieg, and J. K. Gimzewski, Nanotechnology **26**, 204003 (2015).
21) E. C. Demis, R. Aguilera, K. Scharnhorst, M. Aono, A. Z. Stieg, and J. K. Gimzewski, Jpn. J. Appl. Phys. **55**, 1102B2 (2016).
22) D. T. Schoen, C. Xie, and Y. Cui, J. Am. Chem. Soc. **129**, 4116 (2007).
23) M. Ferhat and J. Nagao, J. Appl. Phys. **88**, 813 (2000).
24) A. Sahu, D. Braga, O. Waser, M. S. Kang, D. Deng, and D. J. Norris, Nano Lett. **14**, 115 (2014).
25) T. Kotooka, S. Lilak, A. Stieg, J. Gimzewski, N. Sugiyama, Y. Tanaka, H. Tamukoh, Y. Usami, and H. Tanaka, (doi:10.21203/RS.3.RS-322405/V1).
26) M. Khan, H. Huttunen, and J. Boutellier, Eur. Signal Process. Conf. 2018, p. 682.
27) Q. Wan and X. Fu, Proc. Annu. Int. Symp. Microarchitecture, MICRO, 2020, p. 229.
28) J. H. Shin and T. H. Kim, IEICE Trans. Inf. Syst. **E103.D**, 706 (2020).
29) M. Rastegari, V. Ordonez, J. Redmon, and A. Farhadi, European Conf. Computer Vision 2016 (Springer) 2016, pp. 525–542.

30) K. J. Piczak, Proc. 23rd ACM Int. Conf. Multimedia, 2015, pp. 1015–1018.
31) Y. Yoshimoto and H. Tamukoh, J. Robot. Mechatronics **33**, 386 (2021).
32) H. Yonekawa and H. Nakahara, Proc.-2017 IEEE 31st Int. Parallel Distrib. Process. Symp. Work. IPDPSW 2017 98 (2017).
33) H. Nakahara, T. Fujii, and S. Sato, 2017 27th Int. Conf. F. Program. Log. Appl. FPL 2017, 10.23919/FPL.2017.8056771.
34) Z. Jackson, C. Souza, J. Flaks, and H. Nicolas, (2020), Jakobovski/free-spokendigitdataset:v1.0.10 Available: https://github.com/Jakobovski/free-spoken-digit-dataset.
35) B. Gates, B. Mayers, A. Grossman, and Y. Xia, Adv. Mater. **14**, 1749 (2002).
36) J. Zhang, Q. Fu, Z. Cui, and Y. Xue, CrystEngComm **21**, 430 (2019).
37) B. K. Iwana and S. Uchida, PLoS One **16**, e0254841 (2021).