scientific reports

OPEN

Run-off election-based decision method for the training and inference process in an artificial neural network

Jingon Jang✉, Seonghoon Jang, Sanghyeon Choi & Gunuk Wang✉

Generally, the decision rule for classifying unstructured data in an artificial neural network system depends on the sequence results of an activation function determined by vector–matrix multiplication between the input bias signal and the analog synaptic weight quantity of each node in a matrix array. Although a sequence-based decision rule can efficiently extract a common feature in a large data set in a short time, it can occasionally fail to classify similar species because it does not intrinsically consider other quantitative configurations of the activation function that affect the synaptic weight update. In this work, we implemented a simple run-off election-based decision rule via an additional filter evaluation to mitigate the confusion from proximity of output activation functions, enabling the improved training and inference performance of artificial neural network system. Using the filter evaluation selected via the difference among common features of classified images, the recognition accuracy achieved for three types of shoe image data sets reached ~ 82.03%, outperforming the maximum accuracy of ~ 79.23% obtained via the sequence-based decision rule in a fully connected single layer network. This training algorithm with an independent filter can precisely supply the output class in the decision step of the fully connected network.

The artificial neural network (ANN), which consists of a synaptic device unit capable of both processing and memory functions, has attracted great interest as a brain-inspired analog computing architecture that can efficiently address unstructured complex tasks[1–9]. Particularly, the ANN-based computing system for big data processing covers various application ranges, such as for image[10,11] or voice[12] recognition, signal spectrum analysis[13,14], convolution filtering[15] and use in visual perception systems[16,17]. The conductance modulation capability of a constitutive artificial synaptic node in the ANN can reconfigure the synaptic weight matrix under a learning rule, which is analogous to the essentials of neural plasticity in the human brain. The cumulative output signal is dependent on vector–matrix multiplication using the synaptic weights on each node, which can imitate integrated signal firing through the dendrite of the postneuron in the cortex neural network[18–25]. In this ANN system, the typical decision rule is determined by the sequence of activation function values at each output neuron confined in a fixed range via a monotonic increasing function, such as a sigmoidal or tangent hyperbolic function[26–29]. Based on the magnitude of the fired output values, the classification decision is simply assigned to a certain output neuron that has the highest value. However, this sequence-based decision process is fundamentally limited in its ability to classify data sets in the same object category because it does not consider other configurations of the activation function values. Specifically, when the difference in the activation function values between output neurons is relatively small and the sequence result is incorrectly ordered during training, the ANN system can guide the synaptic weights to the wrong updating direction in a given matrix according to the backpropagation learning rule[30,31]. As a suggested approach for mitigating the abovementioned issue, construction of an auxiliary network in a subsection of the ANN could become an alternative approach that can evaluate the answer among predetermined competitive output neurons in a specified network region at a particular configuration of activation function values. In other words, this approach is based on the run-off election based decision process and is designed to choose a more precise answer between species that possess high scores after the preliminary evaluation.

In this work, we suggest a mutually independent auxiliary synaptic network that can deliver a highly reliable shoe image recognition process in the ANN system by preventing misinterpretations resulting from the

KU-KIST Graduate School of Converging Science and Technology, Korea University, 145, Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea. ✉email: jangjg@korea.ac.kr; gunukwang@korea.ac.kr

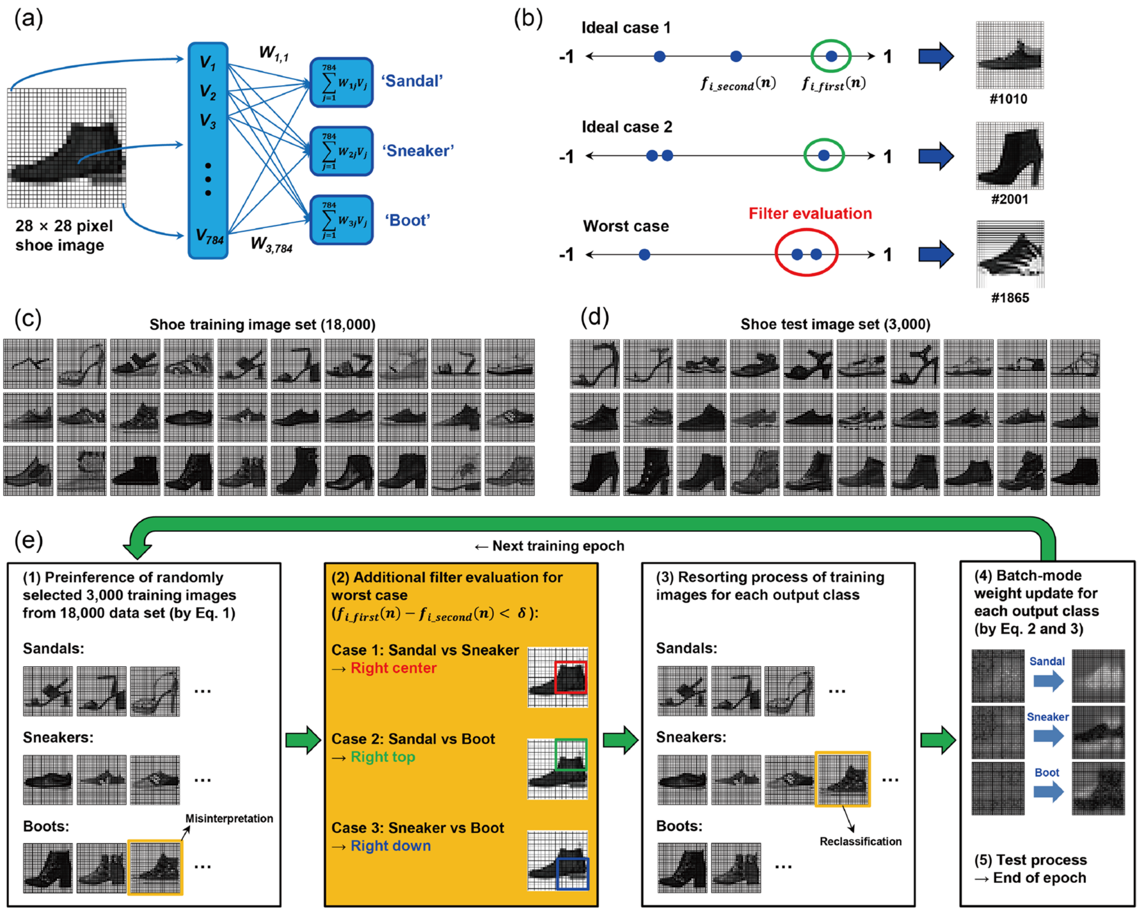

(a)
28 × 28 pixel shoe image
V_1
V_2
V_3
V_{784}
$W_{1,1}$
$W_{3,784}$
$\sum_{j=1}^{784} W_{1j}V_j$
$\sum_{j=1}^{784} W_{2j}V_j$
$\sum_{j=1}^{784} W_{3j}V_j$
'Sandal'
'Sneaker'
'Boot'
(b)
Ideal case 1
-1
1
$f_{i\_second}(n)$
$f_{i\_first}(n)$
#1010
Ideal case 2
#2001
Worst case
Filter evaluation
#1865
(c)
Shoe training image set (18,000)
(d)
Shoe test image set (3,000)
(e)
← Next training epoch
(1) Preinference of randomly selected 3,000 training images from 18,000 data set (by Eq. 1)
Sandals:
Sneakers:
Boots:
Misinterpretation
(2) Additional filter evaluation for worst case
$(f_{i\_first}(n) - f_{i\_second}(n) < \delta)$:
Case 1: Sandal vs Sneaker
→ Right center
Case 2: Sandal vs Boot
→ Right top
Case 3: Sneaker vs Boot
→ Right down
(3) Resorting process of training images for each output class
Sandals:
Sneakers:
Boots:
Reclassification
(4) Batch-mode weight update for each output class (by Eq. 2 and 3)
Sandal
Sneaker
Boot
(5) Test process
→ End of epoch

Figure 1. (**a**) Constituents of FCSN for the 28 × 28 shoe image inference process. Each input pixel is linked to the output neurons consisting of 'Sandal', 'Sneaker' and 'Boot' with a synaptic weight. (**b**) Configuration of the activation function showing 3 output results for each input image (between − 1 and 1) for the ideal cases (upper and middle line) and worst case (lower line). (**c**) 18,000 training images and (**d**) 3,000 test images for the shoe data set. (**e**) Flow chart of one epoch of the training/test algorithm, where the yellow box shows that the filter evaluation is only executed if the difference between the first and second activation function values is smaller than the predetermined δ-value.

conventional decision rule that considers only the sequence of output activation function values. For classification of confusable images, the auxiliary network, which is a quarter of the size of the original pixel image, is partially processed at the assigned filter location selected according to the difference of common features at the preliminary evaluation. In this process, the confusable images are properly reclassified according to the result of the filter evaluation. In other words, the resorting process with the 3-filter auxiliary networks can apply the change in the common feature of the classified output classes and modulation of synaptic weight for the fully connected single layer network (FCSN) which can be determining step in various neural network, leading to a more precise learning direction. As a result, this method can achieve meaningful accuracy improvement for difficult shoe image recognition. This result suggests that filter evaluation based on the run-off election method in the determining step of the ANN system can improve the efficiency of the complex image classification process.

Methods

Conventional neural network for the 3-class shoe image inference process.

The main task is shoe classification solved by vector–matrix calculation of the analogue state level in the FCSN system, as represented in Fig. 1a. Each shoe image consists of 784 (28 × 28) pixels with 256 levels of grayscale[32], which are transformed to the magnitude of the input bias between − 0.5 and 0.5 V. The input neuron is fully connected to a 3-class output neuron, in terms of 'Sandal', 'Sneaker' and 'Boot', with individual connecting strengths as the synaptic weight. The conventional decision method is based on a sequence of vector–matrix multiplication results between the input bias and the selected synaptic weight that is transformed to non-dimensional parameters via the tangent hyperbolic function to limit the output signal in the range from − 1 to 1 as follows[31,33,34]:

$$f_i(n) = \tanh(\beta \sum w_{ij} V_j) \quad (1)$$

where β is the coefficient used to determine the slope of the activation function, w_{ij} is the weight matrix element, and V_j is the pixel intensity of input image (j = 1–784). From this decision rule, a number of image classifications are quickly available in the FCSN system, but an intrinsic limitation exists because the overall arrangement of the output results is not of concern in this step. Indeed, the activation function values of the undesignated output neurons are quite important because they also affect the weight increment process in a negative direction and modify the synaptic weight matrix according to the classification results at each training epoch[30]. If the second highest output neuron is the targeted answer and first output neuron is the inferred answer in confusable classification, this situation inevitably causes a considerably incorrect weight update of the correct image in the backpropagation learning rule[30,31]. The specific example is represented in Fig. 1b with the configuration of activation function values for a few shoe images obtained by simulation in a FCSN system. For the ideal case 1, the output result is well distributed in an overall range, and the first highest value clearly corresponds to 'Sneaker' (1010th image: − 0.88165 for Sandal, − 0.57557 for Boot, and − 0.38362 for Sneaker). Similarly, the decision for the ideal case 2 is also obvious (2001th image: − 0.87317 for Sandal, − 0.81738 for Sneaker and 0.2716 for Boot) because the highest activation function value for the output neuron of 'Boot' is located far away from the other values. However, in the worst case (1865th image: − 0.9034 for Sandal, − 0.50689 for Sneaker and − 0.46234 for Boot), the first and second highest activation function values are so similar that classification is notably difficult between 'Sneaker' and 'Boot'. To overcome this confusion, we established an additional auxiliary network that links a portion of the image pixel as an input bias to the competitive output neurons, and it works only in a particular case when the differences of the first and second highest activation function values are too small to distinguish, i.e., less than the designated limit value.

Run-off election-based simulation algorithm. For implementation of the run-off election decision method in the FCSN system, first, we preliminarily classified the 3,000 training shoe images that were randomly selected from the 18,000-image data set (Fig. 1c) using the conventional decision rule. Because random selection of the training image was different in each training epoch, the reliability of simulation can be enhanced with reflecting complex data variability[35]. In this work, we applied an additional evaluation step to perform a resorting process for the confusable image that shows a smaller difference between the first and second activation function values than the predetermined δ-value by filter evaluation according to the case of competitive classes. For example, if there was confusion between 'Sandal' and 'Sneaker', resorting was performed by vector–matrix multiplication between the input bias in the 196 pixels of the right center (RC) and a 2-class output neuron, similarly for the case of 'Sandal' and 'Boot' in the right top (RT) and 'Sneaker' and 'Boot' in the right down (RD). The classified results, including the results of filter evaluation, were transferred to calculation of the weight increment for each synaptic network using the conventional batch-mode delta rule weight updating process, as shown in the following Eqs. [33,34]:

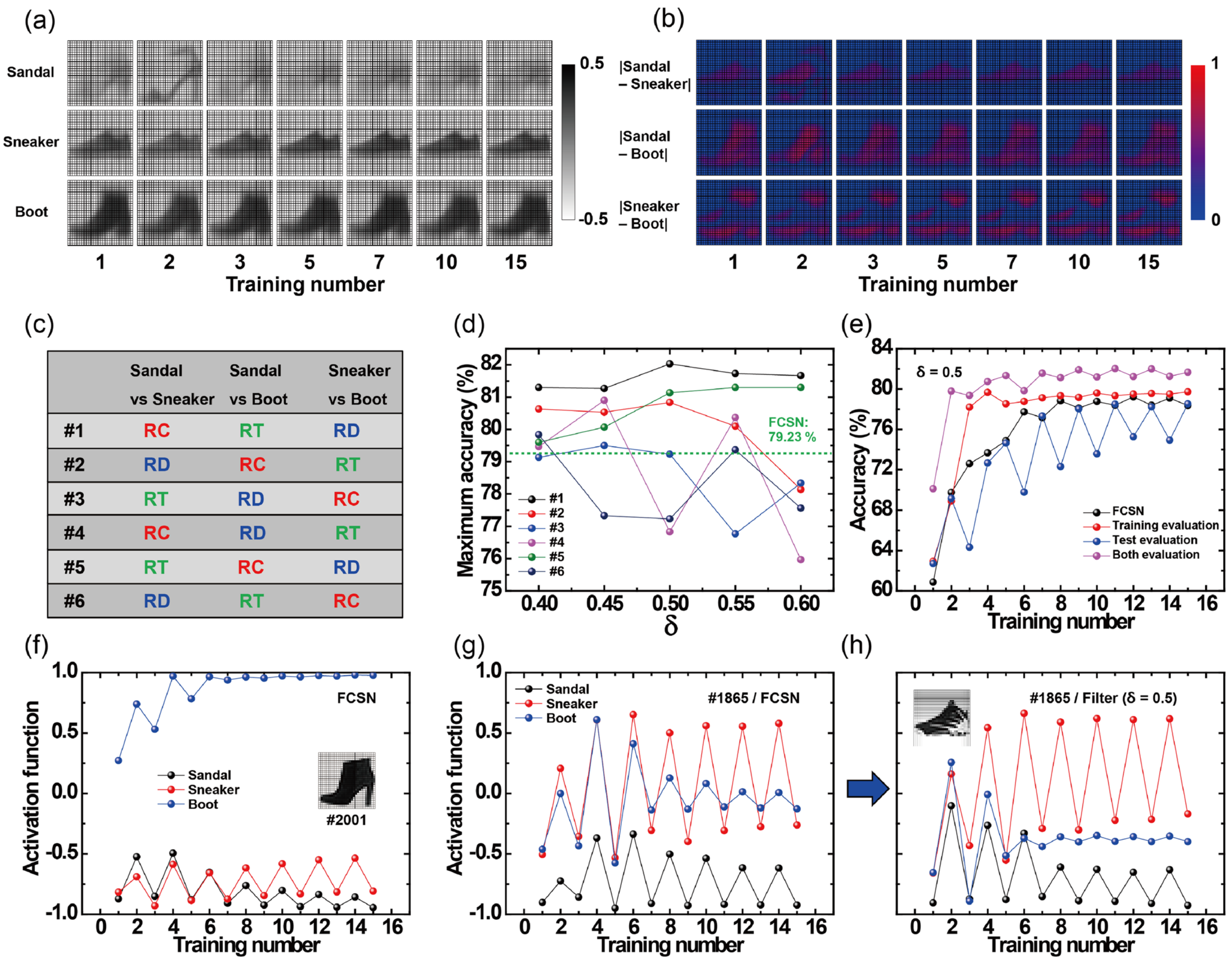
$$\Delta_{\mathrm{ij}}(\mathrm{n}) \;=\; [f_{\mathrm{i}}^{(\mathrm{g})}(\mathrm{n}) - f_{\mathrm{i}}(\mathrm{n})]\; \mathrm{d}f/\mathrm{dI}\; \mathrm{V}_{\mathrm{j}}(\mathrm{n}) \tag{2}$$

$$\Delta \mathrm{w}_{\mathrm{ij}} = \alpha \sum \Delta_{\mathrm{ij}}(\mathrm{n}) \tag{3}$$

where $f_{\mathrm{i}}^{(\mathrm{g})}(\mathrm{n})$ is the target value of the *i*-th output neuron for the *n*-th input image, 'I' means $\Sigma \mathrm{w}_{\mathrm{ij}}\mathrm{V}_{\mathrm{j}}$ (Eq. 1), $\Delta\mathrm{w}_{\mathrm{ij}}$ is the updated weight value for *j*-th pixel of *i*-th output, and α is the device parameter determined by the conductance change at the given conductance value which is obtained by fitting the long-term potentiation/depression (LTP/LTD) curves of the synaptic device because each synaptic weight element corresponds to conductance in synaptic device cell (Supplementary Table 1). Among the various pulse configurations used to obtain the LTP/LTD curves, we focused on the specified condition at a pulse magnitude of − 1.65 V and width of 100 ms. The simulation results for different LTP/LTD curves are presented in Supplementary Fig. 1. The weight updating processes for the 3 types of auxiliary network were generally identical to that of the main network, but they differed in the number of pixels (196) and output (2) elements. Subsequently, 3,000 test images (Fig. 1d) that were not used in the training process were evaluated by ones based on the updated synaptic weight matrix after each training epoch. Figure 1e summarizes one epoch of the run-off election decision algorithm using filter evaluation (yellow box) that implements the resorting process via the 2 × 196 sized auxiliary network according to the competitive classes.

Results and discussion

Effect of filter evaluation for the inference process. Figure 2a shows the overlapped intensity of the classified training images in FCSN for each output class during the training process. Based on this common feature, we obtained the absolute difference contour map between the two output classes (nC_2 cases for n of total output) as shown in Fig. 2b, where the red pixel region shows the more different portion to be selected as filter location while the blue pixel indicates the similar portion. For competition between Sandal and Sneaker, the RC portion (rear end of shoes) showed the most distinguishable intensity, i.e., the appropriate location for the filter evaluation. On the same principle, the RT (ankle portion) and RD (heel portion) region were chosen as an auxiliary filter location for Sandal-Boot and Sneaker-Boot evaluation, respectively. To confirm the availability of this selection, we investigated the filter-evaluated simulation for other configurations of filter location in the RC, RT, and RD regions, as shown in Fig. 2c. The abovementioned filter arrangement is relevant to case #1, which presented the highest accuracy improvement of over ~ 82.03% (black line in Fig. 2d) at a δ-value of 0.5 compared with the original accuracy value of ~ 79.23% in the FCSN (green dash line). The inference results showed variability along the configuration of filter location (e.g., case #4 exhibited a considerable decrease in accuracy), which implied the importance of the proper assignment of filter location along the competitive classes. The statistical results for the filter-evaluated simulation are presented in Supplementary Fig. 2.



	Sandal vs Sneaker	Sandal vs Boot	Sneaker vs Boot
#1	RC	RT	RD
#2	RD	RC	RT
#3	RT	RD	RC
#4	RC	RD	RT
#5	RT	RC	RD
#6	RD	RT	RC

Figure 2. (**a**) Average magnitude of the classified training images for each output class at each training epoch in FCSN and (**b**) absolute difference between the two classes, where the red pixel indicates the more distinguishable region and the blue pixel indicates a similar region. (**c**) Arrangement of the 3-filter location along the competitive output classes. (**d**) The maximum inference accuracy in 15 training numbers depends on the filter arrangement in (**c**), where the #1 case (black line) showed the best improvement at a δ-value of 0.5 compared with the result in the FCSN without filter evaluation (green dash line). (**e**) Evolution of inference accuracy during 15 training numbers for the FCSN (black line) and with filter evaluation only in the training process (red line), only in the test process (blue line), and in both the training/test processes (magenta line). Evolution of activation function values during 15 training numbers for test image (**f**) #2001 and (**g**) #1865 in the FCSN. (**h**) Enhanced separation of activation function values for test image #1865 with filter evaluation at a δ-value of 0.5.

Figure 2e shows comparison of the inference results at different filter application ranges for a δ-value of 0.5. The result for the FCSN is plotted using a black line, which corresponds to a convergence of ~ 79.23% (green dash line in Fig. 2d). When the filter evaluation was partially applied in either the training (red line) or the test (blue line) step, no significant improvement was observed because application in the test process only could not derive the weight matrix change during the training epoch, and only the training process could not utilize the filter decision method. The result of filter evaluation in both the training/test process is represented by a magenta line, which presented remarkably enhanced performance with a maximum accuracy of ~ 82.03%. To assess the propagation of the activation function values in the effect of filter evaluation, we analyzed the evolution of the activation function for the test images relevant to the ideal and worst cases (2001th and 1865th images in Fig. 1b) during the training epoch. Figure 2f shows the result for the ideal case classified as 'Boot' for the 2001th test image in the FCSN. The activation function value for the output neuron of 'Boot' (blue line) shows clear superiority compared with the other output results from the beginning, and its separation was also gradually increased. In this ideal case, filter evaluation is not required because of a wide gap between the first and second highest output results. However, when the first and second output results were sufficiently similar, as in the 1865th image, the classification task was notably difficult because of a continuous incorrect weight update from misjudgment of these images. Indeed, for this case, the confusion of the output result between 'Sneaker' and 'Boot' persisted

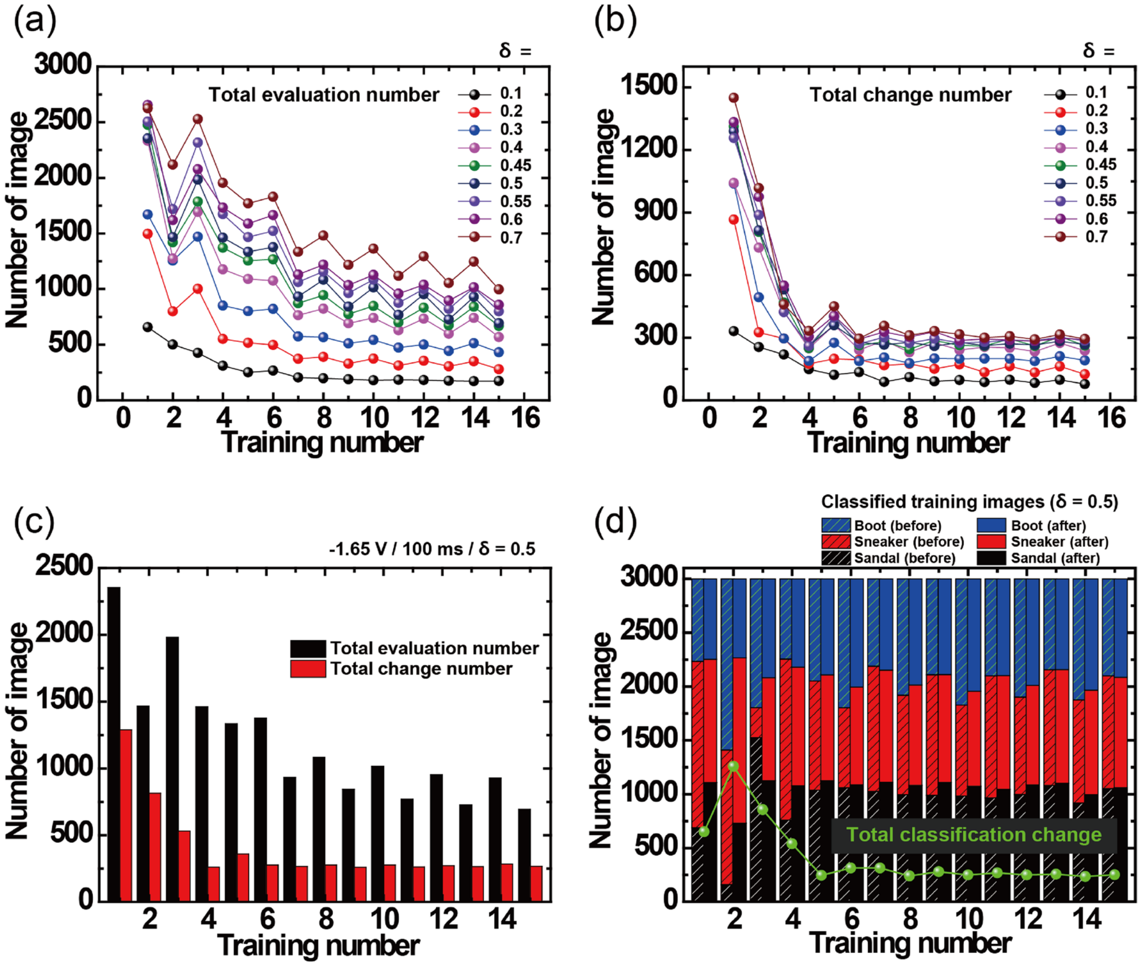


Figure 3. Number of (**a**) total evaluations and (**b**) changes for test images during 15 training numbers with δ-value variation from 0.1 to 0.7. (**c**) Total number of filter-evaluated test images (black bar) and classification changes (red bar) during 15 training numbers at a δ-value of 0.5. (**d**) Classified numbers of training images before (solidus bar) and after (filled bar) filter evaluation during 15 training numbers at a δ-value of 0.5, where the green line represents the number of total classification changes at each epoch.

during the entire training epoch in the FCSN (Fig. 2g). In such a worst case, filter evaluation can be usefully applied to separate the competitive output class. Because the output result for the 1865th image shows confusion between 'Sneaker' and 'Boot', the filter evaluation focused on the RD filter (#1 condition in Fig. 2c) with a δ-value of 0.5. The inference result for the filter evaluation is presented in Fig. 2h, where clear separation of competitive output classes can be observed as the training number increases (red and blue line), producing a reliable decision for 'Sneaker' with the aid of a modified weight matrix update from the resorting process of confusable images.

In the run-off based decision method, the number of evaluation runs is entirely dependent on the conditional constraint of the δ-value because it eventually determines whether the filter evaluation is used at a given difference value between the first and second activation functions. Consequently, it is evident that the number of evaluations could be increased at a larger δ-value. However, a high evaluation number does not always mean greater enhancement of inference accuracy because excessive evaluation definitely contains images that do not need to be evaluated. Figure 3a,b shows the number of total evaluations and changes for the test images during the filter-evaluated training process with δ-value variation. Here, the decreasing tendency of the evaluation number during training process is particular point of conditionally working run-off election-based decision method unlike the consistent filter operation in conventional convolution network system[15,36]. Although both graphs showed an increasing tendency for higher δ-value, the variation in the total change number was not significant compared with the remarkable increase in evaluation number, which indicates that an excessively high δ-value induced unnecessary evaluation and misjudgment. Actually, the simulation results for filter evaluation presented an optimized δ-value near 0.5 to obtain the maximum inference accuracy (Supplementary Fig. 1c). Figure 3c summarizes the result for a δ-value of 0.5, where it can be easily observed that the total change number (red bar) is relatively invariant to the total evaluation number (black bar). This result indicates that the application of filter evaluation should be confined only to confusable images. The classification changes for the training image are presented in Fig. 3d, where the left solidus and right filled bars at each training number indicate the distribution of classified images before and after filter evaluation at a δ-value of 0.5, respectively. In this work, because 3000 training images were randomly selected in a set of 18,000 images at each training number, the

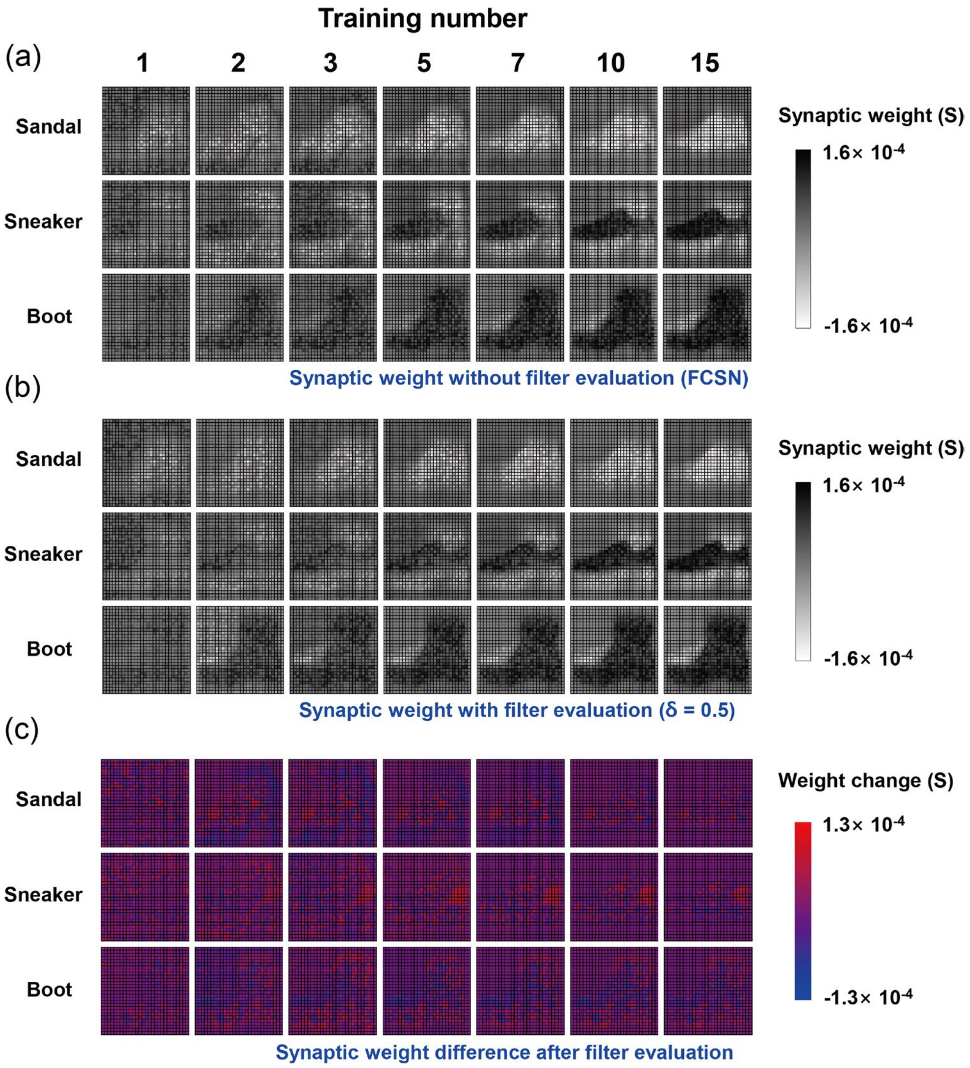


Figure 4. Reshaped 28 × 28 contour images of the synaptic weight for each output class during 15 training numbers (**a**) without and (**b**) with filter evaluation at a δ-value of 0.5 and (**c**) their difference.

training images differed for each epoch, reflecting complex data variability for the simulation. The number of total classification changes are shown by the green line, which indicates that most of the changes were concentrated in the initial training epoch before stabilization of the weight matrix. These results are also related to the notable increase in inference accuracy with filter evaluation, especially at low training numbers (Fig. 2e). Because the sufficiently updated weight matrix become insensitive to additional training steps after a repetitive training process, the correct formation of the weight matrix configuration in the initial step is particularly important[33]. The simulation results for the sigmoidal activation function and without the device parameters are presented in Supplementary Fig. 3.

Synaptic weight updating procedure in the main and auxiliary network. Figure 4a presents the variation of the synaptic weight for each output neuron as the training numbers increase in the FCSN system. Because we used a pair of relevant matrices at each output neuron to link the conductance to the synaptic weight, it was confined in a range of ± $(G_{max} - G_{min})$[33,34]. As the training process continued, the contour map of the weight matrix gradually converged to the shape of the relevant output image because the commonly positioned pixels at each output result received positive weight updates, whereas those of the emptied pixels received negative weight updates according to the learning rule[30,31]. For example, the positioned pixel of 'Boot' enhanced the synaptic weight at the relevant output neuron (black pixel region in the Boot weight matrix) with the training process, but it worked in the negative direction at other output neurons (white pixel region in the Sandal and Sneaker output neurons). Additionally, when the filter evaluation was performed with the shoe resorting process, the propagation of the weight update showed some difference compared with the result in the FCSN. Figure 4b shows the synaptic weight with filter evaluation at a δ-value of 0.5, and the difference value is replotted in Fig. 4c, where a certain portion of the weight element (for example, rear end of Sneaker) was remarkably changed by

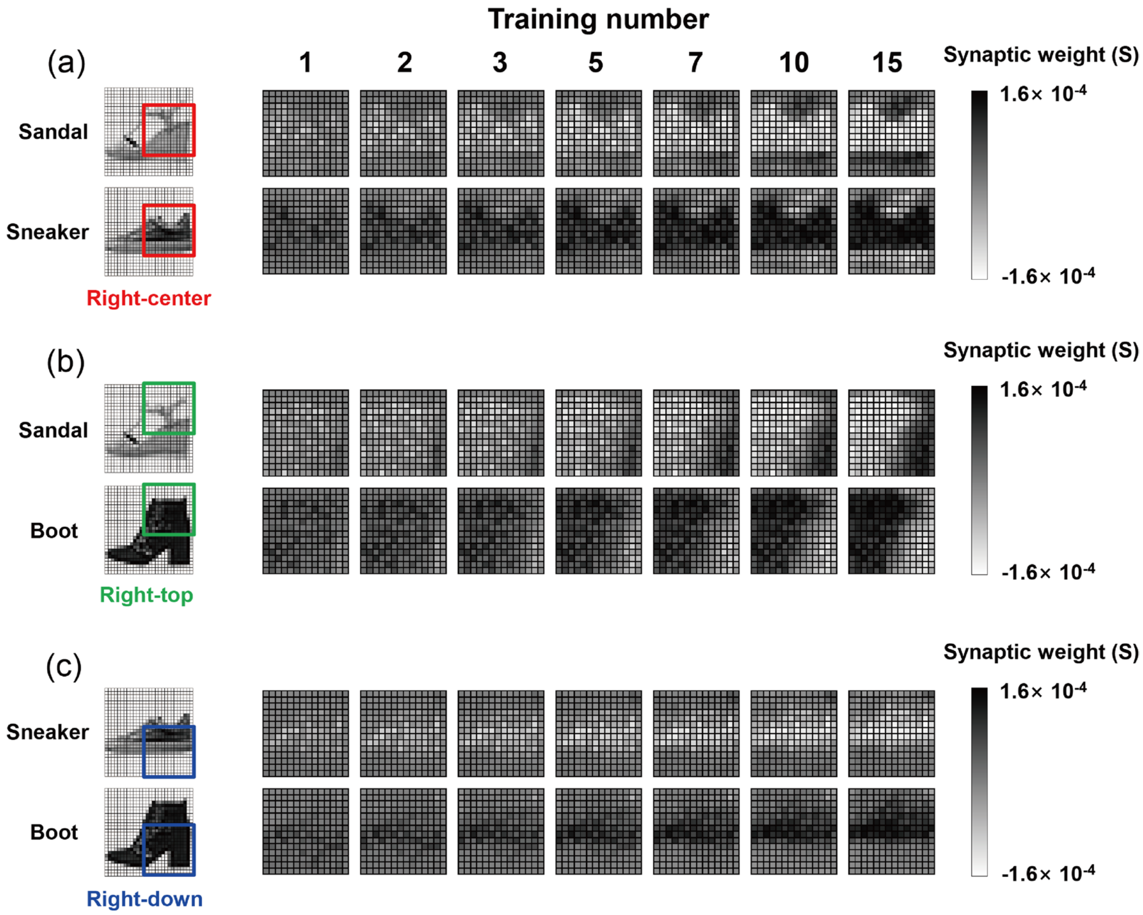


Figure 5. Reshaped 14 × 14 contour images of the synaptic weight of each filter with filter evaluation at a δ-value of 0.5 during 15 training numbers for (**a**) right-center filter to divide 'Sandal' and 'Sneaker', (**b**) right-top filter to divide 'Sandal' and 'Boot', and (**c**) right-down filter to divide 'Sneaker' and 'Boot'.

the effect of classification change in the training process. This detail can also be observed in the synaptic weight of the auxiliary network, as shown in the RC filter (determination of Sandal or Sneaker) in Fig. 5a, where the synaptic weight of the rear end for 'Sneaker' was intensively enhanced via filter evaluation. An interesting point is that the positive weight update quantity in the particular neuron directly affected the other neuron as a result of the negative weight update because of the 2-class filter evaluation in the auxiliary network. In this manner, the complementary weight update process was achieved in RT (Fig. 5b), and the RD (Fig. 5c) filter correctly guided the direction of the synaptic weight update in the main network system. Because the classification change in the training step ultimately results in a change of synaptic weight propagation in all network architectures, a possible interpretation is that the update of the auxiliary network is related to that of the main network system despite its independent synaptic connection, and it is sufficient to infer the main synaptic weight matrix through configuration of the auxiliary network. The synaptic weight propagation of filter-evaluated simulation without device parameters is given in Supplementary Fig. 4 (main network) and Fig. 5 (auxiliary network). In addition, non-square filter application for the generalized data set is presented in Supplementary Fig. 6. Based on the common feature map of the classified output classes, it was enabled to constitute filter shape and location to efficiently separate the competitive output classes, resulting in dramatic enhancement of inference performance of the complex network system.

Conclusion

In summary, we designed a conditionally working auxiliary network to precisely process confusable image classification in the ANN system. By optimizing the constraint of the δ-value and spatial position of each auxiliary network, we could control the quantity of the filter evaluation process and enhance the inference accuracy for the shoe image data set while avoiding misinterpretation caused by confusable configuration of the activation function values. The run-off election-based decision method offers a novel classification rule that can improve the training and inference performance without interruption of the original decision rule based on vector–matrix multiplication, and provides the availability for generalized data set with the selection of filter shape and location based on the common feature of the classified output classes. This method shows the ability to induce a modified synaptic weight update at each node in an array network with the independent auxiliary network. Because the decision step for the output neuron is essential for all types of ANN systems, we believe that this run-off election-based filter evaluation method can be used in a more complex network architecture with superior applicability and effectiveness.

Received: 7 August 2020; Accepted: 8 December 2020
Published online: 13 January 2021

References

1. Bayat, F. M. *et al.* Implementation of multilayer perceptron network with highly uniform passive memristive crossbar circuits. *Nat. Commun.* **9**, 2331 (2018).
2. Wan, Q., Sharbati, M. T., Erickson, J. R., Du, Y. & Xiong, F. Emerging artifcial synaptic devices for neuromorphic computing. *Adv. Mater. Technol.* **4**, 1900037 (2019).
3. Shi, Y. *et al.* Neuroinspired unsupervised learning and pruning with subquantum CBRAM arrays. *Nat. Commun.* **9**, 5312 (2018).
4. Yang, C.-S. *et al.* All-solid-state synaptic transistor with ultralow conductance for neuromorphic computing. *Adv. Funct. Mater.* **28**, 1804170 (2018).
5. Qian, C. *et al.* Solar-stimulated optoelectronic synapse based on organic heterojunction with linearly potentiated synaptic weight for neuromorphic computing. *Nano Energy* **66**, 104095 (2019).
6. Schmidhuber, J. Deep learning in neural networks: an overview. *Neural Net.* **61**, 85 (2015).
7. Kumarasinghe, K., Kasabov, N. & Taylor, D. Deep learning and deep knowledge representation in spiking neural networks for brain-computer interfaces. *Neural Net.* **121**, 169 (2020).
8. Stoliar, P., Yamada, H., Toyosaki, Y. & Sawa, A. Spike-shape dependence of the spike-timing dependent synaptic plasticity in ferroelectric-tunneljunction synapses. *Sci. Rep.* **9**, 17740 (2019).
9. Wright, C. D., Hosseini, P. & Diosdado, J. Beyond von-Neumann computing with nanoscale phase-change memory devices. *Adv. Funct. Mater.* **23**, 2248 (2013).
10. Park, J., Lee, J. & Jeon, D. A 65-nm neuromorphic image classification processor with energy-efficient training through direct spike-only feedback. *IEEE J. Solid-State Circuits* **55**, 108 (2020).
11. Qiao, N. *et al.* A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128K synapses. *Front. Neurosci.* **9**, 141 (2015).
12. Truong, S. N., Ham, S.-J. & Min, K.-S. Neuromorphic crossbar circuit with nanoscale filamentary-switching binary memristors for speech recognition. *Nanoscale Res. Lett.* **9**, 629 (2014).
13. Aamir, S. A. *et al.* An accelerated LIF neuronal network array for a large-scale mixed-signal neuromorphic architecture. *IEEE Trans. Circuits Syst. I Reg. Papers* **65**, 4299 (2018).
14. Rongala, U. B., Mazzoni, A. & Oddo, C. M. Neuromorphic artificial touch for categorization of naturalistic textures. *IEEE Trans. Neural Netw. Learn. Syst.* **28**, 819 (2017).
15. Li, C. *et al.* Analogue signal and image processing with large memristor crossbars. *Nat. Electron.* **1**, 52 (2018).
16. Wang, H. *et al.* A ferroelectric/electrochemical modulated organic synapse for ultraflexible, artifcial visual-perception system. *Adv. Mater.* **30**, 1803961 (2018).
17. Kwon, S. M. *et al.* Environment-adaptable artifcial visual perception behaviors using a light-adjustable optoelectronic neuromorphic device array. *Adv. Mater.* **31**, 1906433 (2019).
18. Wang, Z., Ambrogio, S., Balatti, S. & Ielmini, D. A 2-transistor/1-resistor artificial synapse capable of communication and stochastic learning in neuromorphic systems. *Front. Neurosci.* **8**, 438 (2015).
19. Li, Y. *et al.* Activity-dependent synaptic plasticity of a chalcogenide electronic synapse for neuromorphic systems. *Sci. Rep.* **4**, 4906 (2014).
20. He, Y. *et al.* Dual-functional long-term plasticity emulated in IGZO-based photoelectric neuromorphic transistors. *IEEE Electron. Device Lett.* **40**, 818 (2019).
21. Lammie, C., Hamilton, T. J., Schaik, A. & Azghadi, M. R. Efficient FPGA implementations of pair and triplet-based STDP for neuromorphic architectures. *IEEE Trans. Circuits Syst. I Reg. Papers* **66**, 1558 (2019).
22. Zhou, F. *et al.* Optoelectronic resistive random access memory for neuromorphic vision sensors. *Nat. Nanotechnol.* **14**, 776 (2019).
23. Ahmed, T. *et al.* Time and rate dependent synaptic learning in neuro-mimicking resistive memories. *Sci. Rep.* **9**, 15404 (2019).
24. Du, N. *et al.* Single pairing spike-timing dependent plasticity in BiFeO3 memristors with a time window of 25 ms to 125 µs. *Front. Neurosci.* **9**, 227 (2015).
25. Jo, S. H. *et al.* Nanoscale memristor device as synapse in neuromorphic systems. *Nano Lett.* **10**, 1297 (2010).
26. Mostafa, H. *et al.* Implementation of a spike-based perceptron learning rule using TiO2−x memristors. *Front. Neurosci.* **9**, 357 (2015).
27. Uzan, H., Sardi, S., Goldental, A., Vardi, R. & Kanter, I. Biological learning curves outperform existing ones in artifcial intelligence algorithms. *Sci. Rep.* **9**, 11558 (2019).
28. Ham, S., Choi, S., Cho, H., Na, S.-I. & Wang, G. Photonic organolead halide perovskite artifcial synapse capable of accelerated learning at low power inspired by dopamine-facilitated synaptic activity. *Adv. Funct. Mater.* **29**, 1806646 (2019).
29. Choi, S. *et al.* A self-rectifying TaOy/nanoporous TaOx memristor synaptic array for learning and energy-efficient neuromorphic systems. *NPG Asia Mater.* **10**, 1097 (2018).
30. Hecht-Nielsen, R. Theory of the Backpropagation Neural Network. *International 1989 Joint Conference on Neural Networks* **1**, 593 (1989).
31. Zamanidoost, E., Bayat, F. M., Strukov, D. & Kataeva, I. Manhattan rule training for memristive crossbar circuit pattern classifiers. *IEEE 9th International Symposium on Intelligent Signal Processing (WISP) Proceedings* 1 (2015).
32. Xiao, H., Rasul, K. & Vollgraf, R. Fashion-MNIST: a novel image dataset for benchmarking machine learning algorithms. *Preprint at* https://arxiv.org/abs/1708.07747 *(2017).*
33. Burr, G. W. *et al.* Experimental demonstration and tolerancing of a large-scale neural network (165000 synapses) using phase-change memory as the synaptic weight element. *IEEE Trans. Electron Devices* **62**, 3498 (2015).
34. Prezioso, M. *et al.* Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature* **521**, 61 (2015).
35. Suri, M., Parmar, V., Singla, A., Malviya, R. & Nair, S. Neuromorphic hardware accelerated adaptive authentication system. *IEEE Symposium Series on Computational Intelligence* 1206 (2015).
36. Gao, L., Chen, P.-Y. & Yu, S. Demonstration of convolution kernel operation on resistive cross-point array. *IEEE Electron. Device Lett.* **37**, 870 (2016).

Acknowledgements

The authors acknowledge financial support from the KU-KIST research fund, a Korea University Grant, the National Research Foundation of Korea (NRF-2019R1A2C2003704 and 2019R1A6A3A01095700), and Basic Science Research Program through the NRF funded by the Ministry of Education (2020R1I1A1A01073059).

Author contributions

J.J. and G.W. conceived the idea, designed the experiments and ran the simulation. S.J. and S.C. contributed to data measurement and analysis.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-020-79452-2.

Correspondence and requests for materials should be addressed to J.J. or G.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.