



Biological neurons act as generalization filters in reservoir computing

Takuma Sumi^{a,b} , Hideaki Yamamoto^{a,c,1} , Yuichi Katori^{d,e}, Koki Ito^{a,c} , Satoshi Moriya^a, Tomohiro Konno^f , Shigeo Sato^{a,c} , and Ayumi Hirano-Iwata^{a,b,c,g}

Edited by Terrence Sejnowski, Salk Institute for Biological Studies, San Diego, CA; received October 7, 2022; accepted April 16, 2023

Reservoir computing is a machine learning paradigm that transforms the transient dynamics of high-dimensional nonlinear systems for processing time-series data. Although the paradigm was initially proposed to model information processing in the mammalian cortex, it remains unclear how the nonrandom network architecture, such as the modular architecture, in the cortex integrates with the biophysics of living neurons to characterize the function of biological neuronal networks (BNNs). Here, we used optogenetics and calcium imaging to record the multicellular responses of cultured BNNs and employed the reservoir computing framework to decode their computational capabilities. Micropatterned substrates were used to embed the modular architecture in the BNNs. We first show that the dynamics of modular BNNs in response to static inputs can be classified with a linear decoder and that the modularity of the BNNs positively correlates with the classification accuracy. We then used a timer task to verify that BNNs possess a short-term memory of several 100 ms and finally show that this property can be exploited for spoken digit classification. Interestingly, BNN-based reservoirs allow categorical learning, wherein a network trained on one dataset can be used to classify separate datasets of the same category. Such classification was not possible when the inputs were directly decoded by a linear decoder, suggesting that BNNs act as a generalization filter to improve reservoir computing performance. Our findings pave the way toward a mechanistic understanding of information representation within BNNs and build future expectations toward the realization of physical reservoir computing systems based on BNNs.

machine learning | biological computing | neuronal networks | bioengineering | optogenetics

The brain is a high-dimensional nonlinear system that exhibits intricate dynamics in response to sensory input and is responsible for essential functions, such as perception and motor control (1–6). These functions are physically implemented in a complex network of biological neurons, which are excitable cells carrying ion channel proteins inserted in a bilayer lipid membrane and are connected via synapses. Inherently, ion channels, cell membranes, and synaptic connections are all sources of biological noise, which add stochasticity to membrane potential (7) and signal transmission (8). Despite stochasticity in biological neurons, the population responses are relatively robust. For example, visual stimulation has been shown to stably recruit robust population activity (9, 10) or sequences of population activity (11, 12) in the mouse visual cortex. Consistent sequential activation of neuronal populations evoked by sensory stimuli has also been observed in the auditory and somatosensory cortex (13). Similarly, motor movement and decision-making have been shown to be associated with robust population dynamics in the motor cortex (14, 15) and posterior parietal cortex (16).

As a conceptual paradigm that links such transient population dynamics of high-dimensional nonlinear systems to information processing, “reservoir computing” has been proposed in machine learning (17–20). The model generally consists of input, reservoir, and output layers, and its main feature is that only the weights between the reservoir and output layers are trained in a supervised manner, while the internal weights of the reservoir layer are fixed (20, 21). The tasks covered by reservoir computing include image classification, speech recognition, time-series prediction, and memory. Although reservoir computing was initially proposed to replace computationally demanding recurrent neural networks, it has been shown to have an affinity with many physical systems that exhibit complex dynamics (22, 23), including photonic circuits (24, 25), spintronic devices (26, 27), and soft/organic materials (28, 29). Cultured neuronal networks on uniform substrates, which exhibit a low-pass filtering property on input signals (30), have also been employed as reservoirs to demonstrate their capabilities in pattern classification (31, 32) and motor control (33).

Significance

The brain is a recurrent network of neuronal cells, which are modeled by, but different from, simplified mathematical representations used in typical artificial neural networks. How does the characteristic connectivity in the recurrent network, such as the modular organization, interact with the biophysical properties of the cells to formulate the dynamics and function of biological neuronal networks? To address this question, we used the framework of reservoir computing to readout the stimulus responses of cultured neuronal networks and evaluated their computational properties in pattern classification and timer tasks. We show that reservoirs of biological neurons filter input signals, which can be classified by a linear decoder and that modularity in the recurrent connectivity facilitate the classification performance.

Author contributions: T.S., H.Y., and Y.K. designed research; T.S., H.Y., and K.I. performed research; H.Y., S.M., T.K., and S.S. contributed new reagents/analytic tools; T.S. analyzed data; S.S. and A.H.-I. supervised research; and T.S., H.Y., and A.H.-I. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: hideaki.yamamoto.e3@tohoku.ac.jp.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2217008120/-/DCSupplemental>.

Published June 12, 2023.

Since the internal weights of the reservoir are fixed during training and inference, its network architecture critically affects the performance of a reservoir computing model. In support of this, multiple theoretical studies have recently shown that the implementation of salient features of neuronal connectivity that characterize the nervous systems of animals, such as the modular structure (34, 35), small-worldness (36, 37), and scale-free degree distribution (37), can be advantageous in reservoir computing models of artificial neural networks. For example, Rodriguez et al. explored the optimal modularity of the reservoir layer in a memory task and showed that the highest short-term memory was achieved at a moderate modularity (34). Kawai et al. adjusted the Watts–Strogatz rewiring probability in the reservoir layer and showed that small-world networks reduced the error of nonlinear time-series prediction (36). Furthermore, Suarez et al. showed that direct mapping of the human brain connectome to the reservoir maximizes the memory capacity of a network and that any manipulation to the brain connectome statistics degrades the reservoir performance (38). Among the several canonical features of the nervous systems, the modular structure is especially important, as it is found both macroscopically at the level of cortical areas (39, 40) and microscopically at the level of neurons and synapses (41, 42). Modular networks are “small-world” by their nature, balancing short path lengths and large clustering coefficients (43), and are also evolutionarily conserved (44). However, it remains unclear how such a network architecture integrates with the biophysics of living neurons, e.g., inherent noise and the filtering property, to characterize the dynamical function of biological neuronal networks (BNNs), due partially to the lack of an appropriate experimental platform.

In this study, we used rat cortical neurons cultured on micropatterned substrates combined with optogenetic tools to establish a simple in vitro model of BNNs to investigate how functional modularity affects the information processing capabilities of BNNs under the reservoir computing framework. We investigated how BNNs transform input signals in a high-dimensional state space, focusing on generalization, or categorization, which is the ability of the brain to identify commonalities in different inputs (45). In dynamical terms, this is associated with consistency, wherein a subtle variation in the input signal is canceled to generate a similar dynamical trajectory (46). We show that functional modularity increases the separation of dynamical trajectories and improves the performance in pattern classification tasks and that BNNs act as generalization filters in such tasks to ease categorical learning. Our work paves the way toward a mechanistic understanding of information representation within BNNs and, simultaneously, builds future expectations toward the realization of physical reservoir computing systems based on BNNs.

Results

Reservoir Computing with Biological Neurons. Our reservoir computing system consists of an input layer, reservoir layer, and output layer. The input layer maps multidimensional time series signals to the reservoir layer and was realized by irradiating patterned red light to neurons expressing red-shifted channelrhodopsin ChrimsonR (Fig. 1*A*). The core of the system, the reservoir layer transforms the input signals $u(t)$ to a high-dimensional latent space and generates neuronal responses $x(t)$. The reservoir in the present work was a micropatterned biological neuronal network (mBNN), created by growing primary rat cortical neurons on an engineered substrate fabricated using microcontact printing (47) or a microfluidic device (48). Micropatterns mimicked the modular architecture of the nervous system of the animals and consisted

of four square modules ($200\ \mu\text{m} \times 200\ \mu\text{m}$), wherein adhered neurons formed dense interconnections, connected by thin lines on which neurites grew to couple neighboring modules (Fig. 1*A*). Spontaneous and evoked activity of mBNN was measured by fluorescent calcium imaging using the calcium indicator Cal-520 AM. Finally, in the output layer, a weight matrix W of a linear decoder was trained to transform $x(t)$ to a desired output signal $y(t)$ for classification. The output layer was implemented offline in a custom Python script.

First, we evaluated the classification performance of the reservoir computing system based on mBNN (“mBNN reservoir”) on static spatial patterns (Fig. 1*B*). The input signal $u(t)$ for this task was a three-dimensional vector, whose element corresponded to each of three rectangular patterns bearing identical areas ($200 \times 167\ \mu\text{m}^2$) but projecting to different locations of the mBNN. The rectangular patterns were irradiated one by one for 100 ms with an interval of 10 s. The task here was to generate a pulse signal of 2.5 s in the output node that corresponded to the input class (top, middle, or bottom). For this, the output weights W were trained by ridge regression to generate the target signal $\hat{y}(t)$ from the relative fluorescence unit (RFU) $\Delta F/F$ of 60 neurons $x(t)$ (Fig. 1*C* and *SI Appendix, Fig. S1A*). The output vector $y(t) = Wx(t)$ was then summed over time for a duration of 2.5 s, and a node with the largest value (argmax) was regarded as the estimated answer from the reservoir system. The accuracy of each mBNN reservoir was evaluated by dividing the number of correct estimates by the number of test trials.

The classification accuracy of the mBNN reservoirs after training was above chance level and significantly higher than the value obtained from a label-shuffled null model, whose accuracy was estimated by assigning random class labels to each of the reservoir responses $[x(t)]$ [two-sided paired t test, $P < 0.001$, $n = 21$, degrees of freedom (df) = 20] (Fig. 1*D*). *SI Appendix, Fig. S1B* summarizes the comparison of the accuracy using either the raw calcium fluorescence, RFU, or their z -scored values as the source of the reservoir state, $x(t)$. Independent of the signal source, the mBNN reservoir could be successfully trained as a spatial pattern classifier (raw data, $71.9 \pm 23.8\%$; z -scored raw data, $72.4 \pm 26.6\%$; RFU, $74.8 \pm 20.2\%$; z -scored RFU, $72.9 \pm 25.3\%$; label-shuffled, $32.4 \pm 13.0\%$; mean \pm SD, $n = 21$ samples).

Furthermore, we calculated the mean squared error (MSE) between $y(t)$ and $\hat{y}(t)$ to evaluate the reservoir performance independent of the argmax operation for classification. The result showed that the MSE of the mBNN reservoirs was significantly smaller than that obtained from a label-shuffled null model (Fig. 1*E*; two-sided paired t test, $P < 0.001$, $n = 21$, $df = 20$), as expected from the evaluation of the classification accuracy. The lowest MSE was obtained when RFU without z -score standardization was used as $x(t)$ (raw data, 0.055 ± 0.022 ; z -scored raw data, 0.104 ± 0.020 ; RFU, 0.047 ± 0.016 ; z -scored RFU, 0.090 ± 0.028 ; label-shuffled, 0.119 ± 0.018 ; mean \pm SD, $n = 21$ samples) (*SI Appendix, Fig. S1C*; two-sided paired t test, $P < 0.001$ (vs. raw data), $P < 0.001$ (vs. z -scored raw data), $P < 0.001$ (vs. z -scored RFU), $n = 21$, $df = 20$). Thus, we used RFU as $x(t)$ in all subsequent experiments.

The classification property depended not only on the source of the reservoir state but also on the nature of the input and readout. To show this, we first examined whether the location of the photostimulation influenced the classification accuracy. Analysis of the confusion matrix (Fig. 1*F*) showed that the lowest false-positives and false-negatives were observed between the top and bottom inputs, which had the largest spatial distance between the irradiated regions. The result indicates that spatially distant inputs were easier to be classified by the mBNN reservoir. Next, we

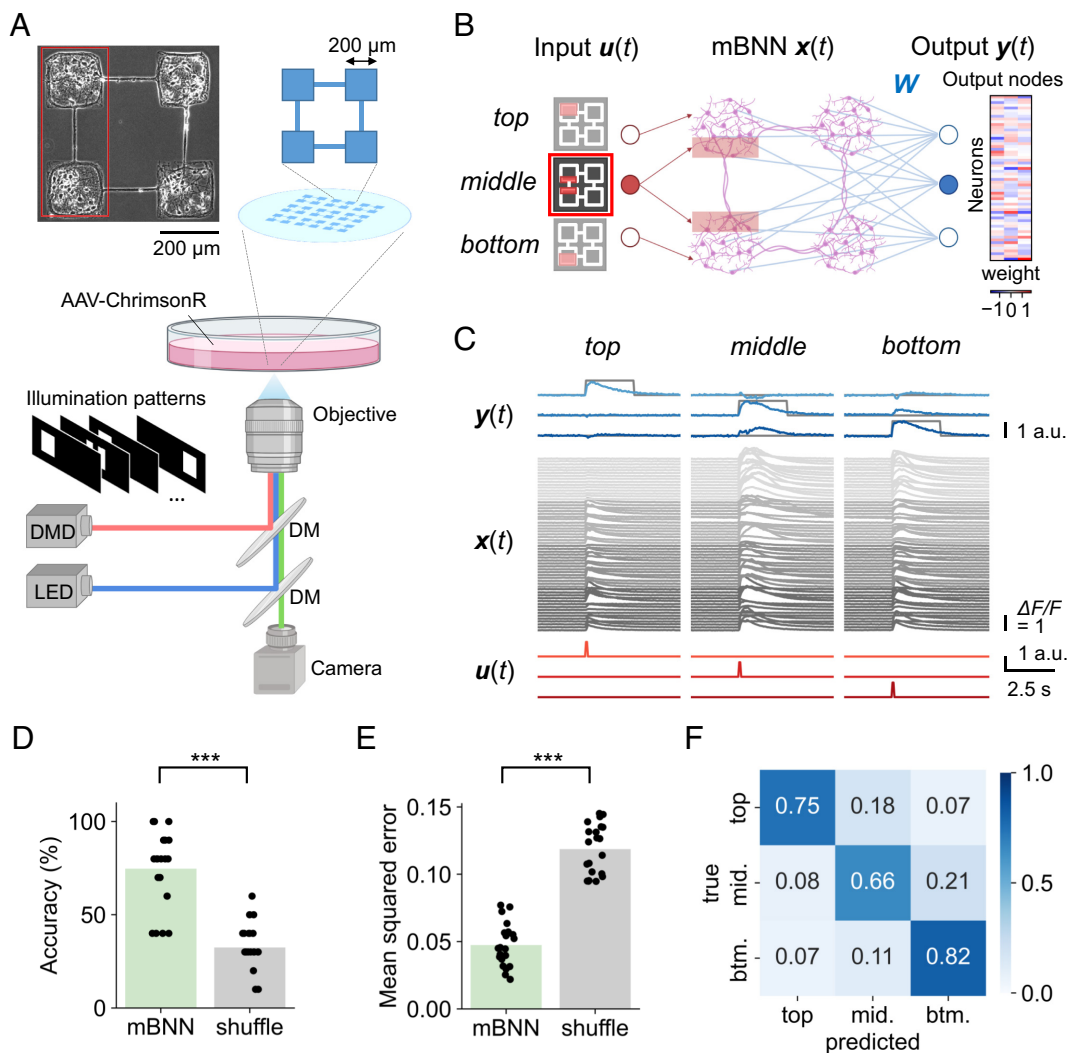


Fig. 1. Reservoir computing based on mBNN. (A) Phase-contrast micrograph of a mBNN and schematic representation of the experimental setup. Photostimulation was delivered to neurons located in the two modules outlined with a red rectangle (*Top*). DMD, digital mirror device; LED, light-emitting diode; DM, dichroic mirror. (B) Overview of the reservoir computing system. The input signal $u(t)$ was a three-dimensional vector, whose elements corresponded to each of the three rectangular patterns (*Top*, *Middle*, *Bottom*). The rectangular patterns of red light were projected to the corresponding regions in an mBNN. The dynamics of 60 neurons (15 neurons per module) in the mBNN $x(t)$ in response to the input was decoded by a linear decoder to generate the output signal $y(t)$. A representative distribution of weight values in the output weight matrix W is shown in the right panel. (C) Representative waveforms of $u(t)$, $x(t)$ (in RFUs), and $y(t)$. Gray traces shown together with $y(t)$ are the target signals $y(t)$ used to train W . (D) Classification accuracy and (E) MSE of the mBNN reservoir computed using the RFU as $x(t)$. “Shuffle” denotes the values obtained using label-shuffled datasets. Filled circles represent a single mBNN, and bar heights show the mean. *** $P < 0.001$ (two-sided paired t test; $n = 21$, $df = 20$). (F) A normalized confusion matrix averaged over mBNNs ($n = 21$). Abbreviations: mid., middle; btm., bottom, which correspond to the input patterns shown in left panels of (B).

investigated the effect of input area on the reservoir performance. The classification accuracy remained constant even when the area of the photostimulation pattern was halved to that of the original pattern. However, a significant decrease was observed when the photostimulation area was quartered (*SI Appendix*, Fig. S2A; two-sided paired t test, $P < 0.05$, $n = 5$, $df = 4$), which most probably is caused by the decreased consistency in the network responses. Finally, in order to investigate the influence of the number of neurons used for readout (N) on the reservoir performance, we varied N from 1 to 60, randomly selected the corresponding number of neurons from the network, and estimated the classification accuracy for each N . Analysis revealed that, the accuracy monotonically increased with N , which saturated by $N = 60$ in most samples (*SI Appendix*, Fig. S2B). The choice of neurons for readout was also crucial. While the classification accuracy for $N = 15$ was 52.5% when the neurons were selected from a single module, the value significantly increased to 57.3%, 59.8%, and 61.5% when the same number of neurons were selected from two,

three, and four modules, respectively (*SI Appendix*, Fig. S2C; two-sided paired t test, $P < 0.001$, 0.001, and 0.001, respectively, $n = 21$, $df = 20$). In summary, given a proper setting of the input and output, mBNN is capable of transforming static input to neuronal dynamics, which can then be readout by a linear decoder for classification.

Modularity Increases the Accuracy of the Spatial Pattern Classification. Next, we investigated how nonrandom connectivity, especially modular organization, impacts the classification performance of mBNN reservoirs. To assess the functional connectivity of mBNNs, we evaluated the correlation coefficient between neurons i and j , r_{ij} , from the recordings of spontaneous activity and calculated the modularity Q of the correlation matrix $[r_{ij}]$ in each network (49). Fig. 2A shows the correlation matrices of two representative mBNNs with high ($Q = 0.13$) and low ($Q = 0.014$) modularity. Notably, the functional connectivity (correlation matrix) of the highly modular network was

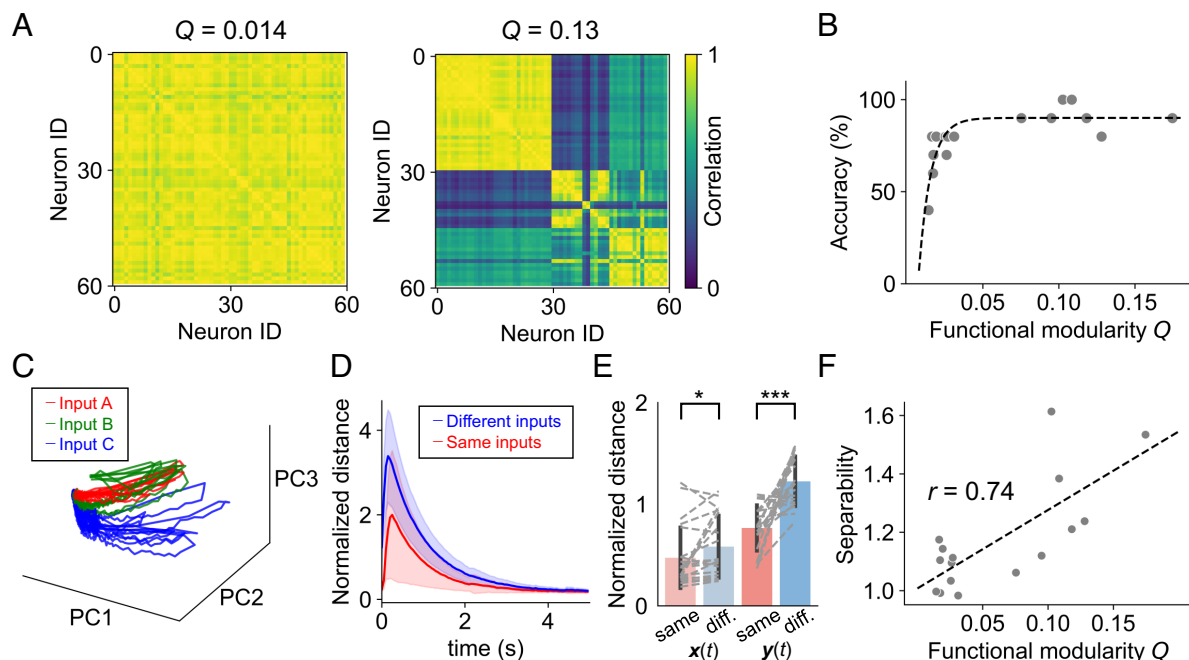


Fig. 2. Structure-function relationships in mBNN reservoirs. (A) Representative correlation matrices of two mBNNs with weak (Left) and strong (Right) modularity. Correlations between two neurons are evaluated using the recordings of spontaneous activity. (B) Classification accuracy in the spatial pattern classification task as a function of functional modularity. Weakly modular networks show large variability in accuracy, whereas strongly modular networks stably show high accuracy. A single exponential curve fit is also shown to aid visualization (dashed line). (C) Trajectories of reservoir dynamics $\mathbf{x}(t)$ visualized in a three-dimensional principal component (PC) subspace. The explained variance ratios are 0.73, 0.23, and 0.01 for PC1, PC2, and PC3, respectively. (D) Instantaneous normalized distance between two trajectories upon receiving the same or different inputs. Mean value at each time are plotted with shades indicating the SD. (E) Comparison of the time-averaged normalized distance between trajectories for the same (D_{same}) and different (D_{diff}) inputs in $\mathbf{x}(t)$ and $\mathbf{y}(t)$. Bar heights show the mean, error bars show the SD, and dotted gray lines show individual mBNN. $*P < 0.05$, $***P < 0.001$ (two-sided paired t test; $n = 21$, $df = 20$). (F) Separability of the trajectories ($D_{\text{diff}}/D_{\text{same}}$) as a function of functional modularity. The Pearson correlation coefficient $r = 0.74$ and $P < 0.001$ (two-sided Pearson correlation test, $n = 16$, $df = 14$).

characterized by the existence of subpopulations of neurons with strong coupling inside the four squares. Such substructures were absent in the functional connectivity of weakly modular networks, whose r_{ij} was ~ 1 for almost all neuron pairs. Analysis of the relationship between functional modularity and computational performance revealed that classification accuracy was highly variable in networks with low modularity (< 0.05), with a trend toward a positive correlation (Fig. 2B). In contrast, for networks with high modularity (> 0.05), the accuracy was mostly independent of modularity, and the classification accuracy was consistently high ($\geq 80\%$). The results suggest that modularity Q reflects meaningful differences in the network performance at least for the lower range of the modularity and that functional modularity within mBNNs is advantageous for the classification of static spatial inputs with a linear readout.

As noted in the previous section (SI Appendix, Fig. S2C), the spatial arrangement of the reservoir neurons used for readout impacts the classification accuracy. Further analysis revealed that the functional modularity was correlated with the ratio of the accuracy when the neurons were selected from four modules against that when the neurons were selected from a single module (SI Appendix, Fig. S2D; two-sided Pearson correlation test, $r = 0.61$, $P < 0.05$, $n = 16$, $df = 14$). This implies that highly modular networks represent information wider in space and that such reservoirs benefit more by reading out from multiple modules.

To understand the mechanisms underlying the improved accuracy in highly modular networks, we analyzed the trajectories of the reservoir dynamics $\mathbf{x}(t)$ in response to each input (Fig. 2C). The similarity between the two trajectories was quantified by calculating their time-averaged normalized distances (Materials and Methods) (Fig. 2D). Our analysis revealed that the mean distance between $\mathbf{x}(t)$

evoked by inputs belonging to the same class (D_{same}) was 0.47 ± 0.30 , while that for different classes (D_{diff}) was 0.58 ± 0.30 and was significantly larger (Fig. 2E; two-sided paired t test, $P < 0.05$, $n = 21$, $df = 20$). Although the difference between the same and different classes was significant but subtle, comparison of the distance in the output signals $\mathbf{y}(t)$ shows that the linear readout by the output layer increases this difference (Fig. 2E). Defining the ratio of D_{diff} to D_{same} as the separability of the reservoir, we found that its mean increased from 1.4 to 1.8 when they are estimated from $\mathbf{y}(t)$ instead of $\mathbf{x}(t)$. Furthermore, investigation of its dependence on the functional modularity showed that separability in each mBNN was significantly correlated with its modularity (Fig. 2F; two-sided Pearson correlation test, $r = 0.74$, $P < 0.001$, $n = 16$, $df = 14$) and accuracy (SI Appendix, Fig. S2E). These results indicate that the increased separability of reservoir dynamics underlies the improved classification performance of highly modular networks.

Performance of mBNN Reservoirs in Speech Recognition. The strength of reservoir computing is its ability to process time-series data using the transient dynamics of the reservoir (19). In pattern recognition tasks, the reservoir must possess “short-term memory” (50) to retain the information of the past input for a finite period. Such a reservoir property can be quantified using the timer task, a simple task that recalls the same signal as the input $u(t)$ with a delay τ (28, 51). As in the previous task, the input waveform was a square pulse with a duration of 100 ms during which one of the rectangular patterns was irradiated to the mBNN. The target signal for the timer task with a delay τ , $\hat{y}_\tau(t)$, was set equal to $u(t - \tau)$.

Fig. 3, Bottom, Inset shows the output signals y_τ from mBNN reservoirs trained to generate \hat{y}_τ with $\tau = 0, 0.4$, and 1 s. As a

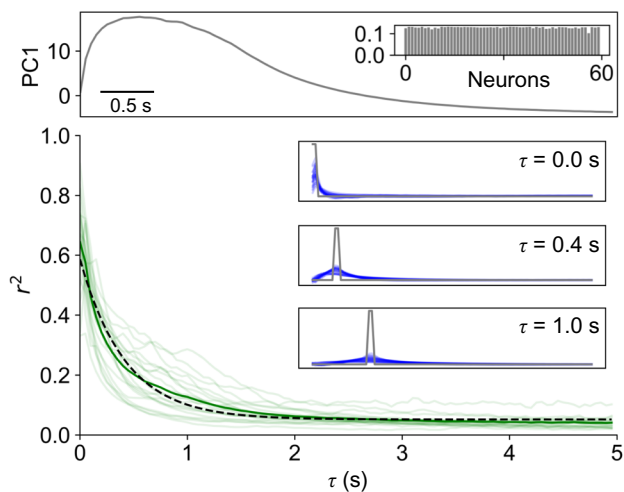


Fig. 3. Memory capacity of mBNNs. *Top* panel, time course of the first principal component (PC1) of the stimulus responses in a representative mBNN (averaged over 30 trials). PC1 represents a network-wide collective activity, as can be seen from the factor loadings in the *Inset*. *Bottom* panel, coefficient of determination r^2 plotted as a function of time delay τ . Light green curves represent a single mBNN, a dark green curve is the mean over all samples ($n = 21$), and a dashed curve is an exponential fit to the mean. *Inset* shows the target function (gray) and mean reservoir outputs of all samples (blue, $n = 21$) at designated time delays. Difference between the reservoir output and the target signal increases with τ .

general trend, the similarity between y_τ from \hat{y}_τ , which could be quantified using the coefficient of determination r^2 , decreased with τ . The dependence of r^2 on τ for all networks ($n = 21$) is

summarized in Fig. 3 and *SI Appendix*, Fig. S3. The r^2 value peaked at 0 to 50 ms and decreased with increasing τ . An exponential fitting the mean value of r^2 across all samples revealed a characteristic decay constant of the curve to be 0.42 s. This result indicates that mBNNs have short-term memory in the order of 100 ms and suggests that they can effectively be used as a reservoir that processes time-series data with a timescale of ~ 1 s.

We thus designed a classification task for a human spoken digit dataset, TI-46, whose signal durations matched the short-term memory of the mBNN reservoir. To deliver spoken digit signals into the reservoir layer, the signal was first converted to a 17-channel cochleagram using Lyon's passive ear model. Photostimulation patterns were then designed by mapping each channel of the cochleagram to individual neurons expressing ChrimsonR and adjusting the spot size of the illumination according to its instantaneous spectral intensity (Fig. 4A). Spoken digit signals of "zero" and "one" were used in this study. The weight matrix of the output layer was then trained using ridge regression to generate a pulse signal corresponding to the input class, and the performance was evaluated following the same procedure as in the spatial pattern classification task.

Visualization of representative reservoir trajectories $\mathbf{x}(t)$ in the principal component (PC) subspace showed that mBNNs generated characteristic responses for the two input classes (Fig. 4B). The mean accuracy in classifying the two digits spoken by a female speaker was found to be $82.5 \pm 21.9\%$ (Fig. 4C), which was significantly higher than the value obtained from label-shuffled null models (two-sided paired t test, $P < 0.001$, $n = 20$, $df = 19$). An equivalent classification accuracy was also obtained even when the number of digits was increased to three, while the value was found

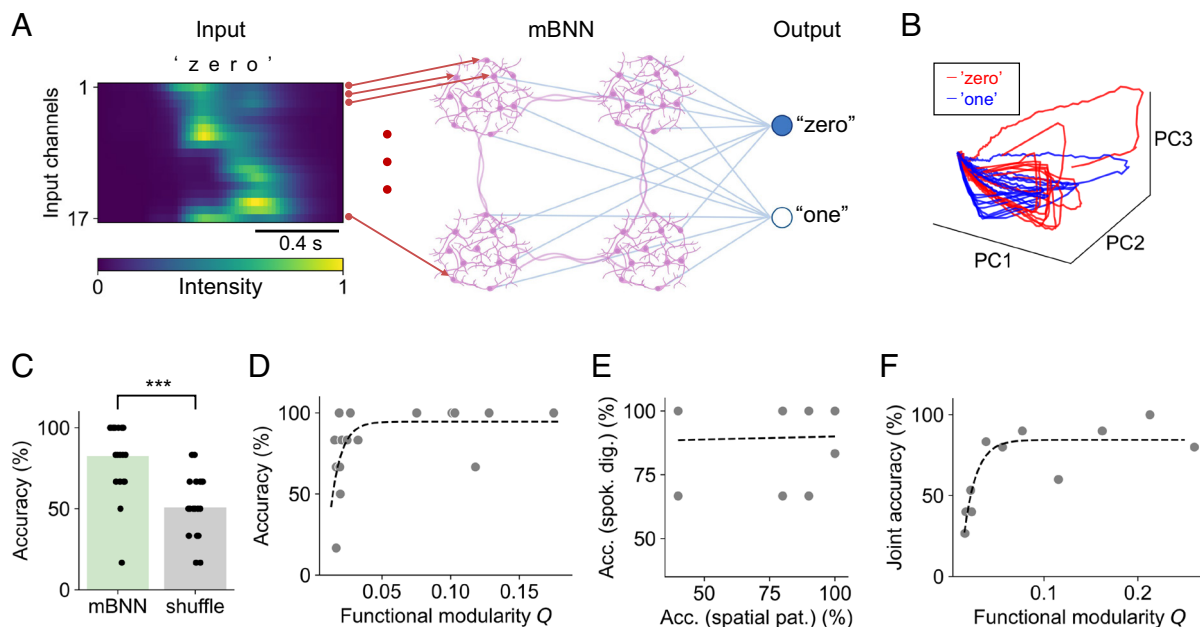


Fig. 4. Classification of the time-series data. (A) Overview of the spoken digit classification task. Audio waveforms are first converted to a 17-channel cochleagram, which are then mapped onto the neurons in mBNNs by assigning one neuron for one frequency band. The input signal is then optogenetically delivered to mBNNs, and the evoked activity is then decoded in the output layer. (B) Trajectories of reservoir dynamics $\mathbf{x}(t)$ visualized in a three-dimensional principal component subspace. The explained variance ratios are 0.90, 0.04, and 0.02 for PC1, PC2, and PC3, respectively. (C) Classification accuracy of the mBNN reservoir. "Shuffle" denotes the values obtained using label-shuffled datasets. Filled circles represent a single mBNN, and bar heights show the mean. The discretization of accuracy values, as well as the low accuracy well below the chance level observed in some samples, is caused by the limited number of test trials ($= 6$). *** $P < 0.001$ (two-sided paired t test, $n = 20$, $df = 19$). (D) Classification accuracy in the spoken digit classification task as a function of functional modularity. As was for the spatial pattern classification task, weakly modular networks show large variability in accuracy, whereas strongly modular networks stably show high accuracy. A single exponential curve fit is shown to aid visualization (dashed line). (E) Accuracy in the spoken digit classification task does not correlate with the accuracy in the spatial pattern classification task. A linear fit is shown to aid visualization (dashed line). The Pearson correlation coefficient $r = 0.04$ and $P = 0.91$ (two-sided Pearson correlation test, $n = 11$, $df = 9$). (F) Joint accuracy, i.e., a product of the accuracies in the spoken digit and spatial pattern classification tasks, as a function of modularity. Note that the functional modularity calculated from evoked activity is used in the abscissa due to the limited number of samples with the recordings of spontaneous activity. The functional modularity calculated from the spontaneous and evoked activities were, however, strongly correlated (*SI Appendix*, Fig. S6; $r = 0.80$ and 0.73). A single exponential curve fit is shown to aid visualization (dashed line).

to decrease significantly in a four-digit classification task (SI Appendix, Fig. S4A; two-sided paired t test, $P < 0.05$, $n = 8$, $df = 7$).

In a control experiment where the classifier was trained to regress directly from the photostimulation patterns illuminated on a silicon wafer (without the mBNN), a classification accuracy of 100% was achieved for the two-digit classification task. Such a classifier, however, failed to distinguish signals having identical spatial components, such as “zero” and temporally reversed “zero” (SI Appendix, Fig. S4B). Importantly, the classifier could distinguish temporally reversed version of a spoken digit from its original when it was trained to regress from the mBNN activity (SI Appendix, Fig. S4C; two-sided paired t test, $P < 0.001$, $n = 9$, $df = 8$). These results indicate that the mBNN reservoir can be used to classify not only static spatial patterns but also complex spatiotemporal patterns, such as human spoken digits.

Although the majority of mBNNs showed a high classification accuracy of over 80%, the accuracy exhibited a large network-to-network variation, ranging from 17 to 100% (Fig. 4C). Thus, to understand the mechanism that supports high-accuracy reservoirs, we assessed the relationship between memory capacity ($= \sum_{\tau} r^2(\tau)$), where $r^2(\tau)$ is the coefficient of determination between y_{τ} and \hat{y}_{τ} and the accuracy of each mBNN. Interestingly, memory capacity did not correlate with accuracy (SI Appendix, Fig. S5; $n = 11$), which we interpret to be caused by the sufficiently high memory capacity of all samples for classifying spoken digit signals. Thus, we analyzed the relationship between functional modularity Q and accuracy (Fig. 4D). Clearly, a large variability in accuracy was observed mostly in weakly modular networks ($Q < 0.05$), whereas the majority ($=5/6$) of strongly modular networks ($Q > 0.05$) exhibited an accuracy of 100%. This trend is consistent with that obtained from the spatial pattern classification task, suggesting that functional modularity is beneficial for both spatial and spatiotemporal pattern classifications.

Surprisingly, a comparison of the accuracy in the spatial pattern and spoken digit classification tasks revealed that the samples that exhibited high performance in one task did not necessarily perform well in the other task (Fig. 4E; two-sided Pearson correlation test, $r = 0.12$, $P = 0.91$, $n = 11$, $df = 9$). Here, again, functional modularity plays an essential role. Fig. 4F shows the joint probability of obtaining correct estimates in the two tasks as a function of modularity, showing that the probability is positively correlated with Q in its lower range. Therefore, mBNN reservoirs with a high Q are stable classifiers independent of tasks.

Generalization Capability of mBNN Reservoirs. Finally, to test the generalization capability of the mBNN reservoir, we designed a generalized spoken digit classification task, wherein the reservoir was first trained to classify the “zero” and “one” inputs of a female or a male speaker and then tested its classification accuracy on “zero” and “one” inputs of a different gender (Fig. 5A). The input spectrum, and hence the resulting photostimulation pattern, of the female/male speakers was distinct, which could be distinguished with an accuracy of 100% by the linear classifier (Fig. 5B). However, when the classifier was trained to regress directly from the photostimulation patterns, the classification accuracy decreased to the chance level (50%) if the speaker was switched in the training and testing phases (Fig. 5B).

In contrast, when the signal was passed through the mBNN, the classification accuracy of $64.1 \pm 15.1\%$ ($n = 16$) was achieved (Fig. 5C). The value was 18.8% lower than that obtained without switching ($82.8 \pm 15.4\%$), but significantly higher than the value obtained from the label-shuffled models (two-sided paired t test, $P < 0.001$, $n = 16$, $df = 15$). These results indicate that the classification is achieved nonrandomly, and the signal transformation

by the mBNN endows the reservoir with a generalization ability to distinguish spoken digits independent of the speakers.

To provide further evidence for the generalization ability of mBNN, we designed a gender classification task, wherein the reservoir was first trained to classify the spoken digit “zero” of a male and a female speaker and then tested to classify the gender of the speaker on the spoken digit “one” (Fig. 5D). Similar to the previous task, without the mBNN, an output layer trained on photostimulation images failed to classify gender when the spoken digits were switched (Fig. 5E). In contrast, the signal transformation by the mBNN increased the classification accuracy to 71.4%, which was well above the chance level. Although the accuracy was lower than that obtained without the switching of the digits, it was higher than that of the label-shuffled model (Fig. 5F; two-sided paired t test, $P < 0.001$, $n = 16$, $df = 15$). These observations support the finding that the mBNN acts as a generalization filter for the spatiotemporal input patterns.

To understand which features of mBNNs enabled generalization, we investigated how each input signal was transformed after passing through the mBNN. The trajectories of photostimulation patterns $\mathbf{u}(t)$ and reservoir layer $\mathbf{x}(t)$ projected onto three-dimensional PC subspaces are shown in Fig. 5G. The trajectories of $\mathbf{u}(t)$ for the same digits spoken by the same speaker were constrained to a narrow tube, whereas those for different digits and speakers were separated. In contrast, the trajectories of $\mathbf{x}(t)$ showed a large variability even for the same input classes but still formed widened tubes, which were separated for different classes. The mean normalized distance of the trajectories for the same class (D_{same}) was found to be 0.07 and 0.56 for $\mathbf{u}(t)$ and $\mathbf{x}(t)$, respectively. For different classes, the value (D_{diff}) was 0.57 and 0.69 for $\mathbf{u}(t)$ and $\mathbf{x}(t)$, respectively (Fig. 5H). This indicates that an mBNN transforms its input to reduce the difference between D_{diff} and D_{same} while retaining the magnitude relation of $D_{\text{diff}} > D_{\text{same}}$. Such a transformation enabled the generalized classification of signals of different genders or digits.

Finally, we quantified how the switching of digits and speakers affects the dynamics in the mBNN by evaluating the following three variables: D_{digits} , the mean distance between the trajectories for the same speakers but different digits; D_{speaker} , that for the same digits but different speakers; and D_{both} , that for different speakers and digits. Distance analysis for $\mathbf{x}(t)$ revealed that D_{both} was significantly greater than both D_{digits} (two-sided paired t test, $P < 0.001$, $n = 16$, $df = 15$) and D_{speaker} (two-sided paired t test, $P < 0.001$, $n = 16$, $df = 15$) (Fig. 5I). Therefore, the separation between different classes was quantitatively retained in $\mathbf{x}(t)$, exhibiting the largest separation when both digits and speakers were different. In contrast, while D_{both} in $\mathbf{u}(t)$ was 0.37 greater than D_{speaker} , the value was 0.01 smaller than D_{digits} . This mingled relationship between the distances is considered to have caused the failure in generalized classification from $\mathbf{u}(t)$. In summary, the mBNN reduces the separation of trajectories between different classes in the high-dimensional state space while preserving the commonalities between different classes, enabling the distinction of different speakers and digits through a simple, linear decoder.

Discussion

In this study, we constructed mBNNs with modular topology on engineered substrates and demonstrated the effectiveness of the modular architecture and biological signal transduction in pattern classification tasks under the reservoir computing framework. Our experiments demonstrated that the accuracy of pattern classification was positively correlated with the functional modularity of BNNs in the lower range of the modularity Q . Conventionally, a

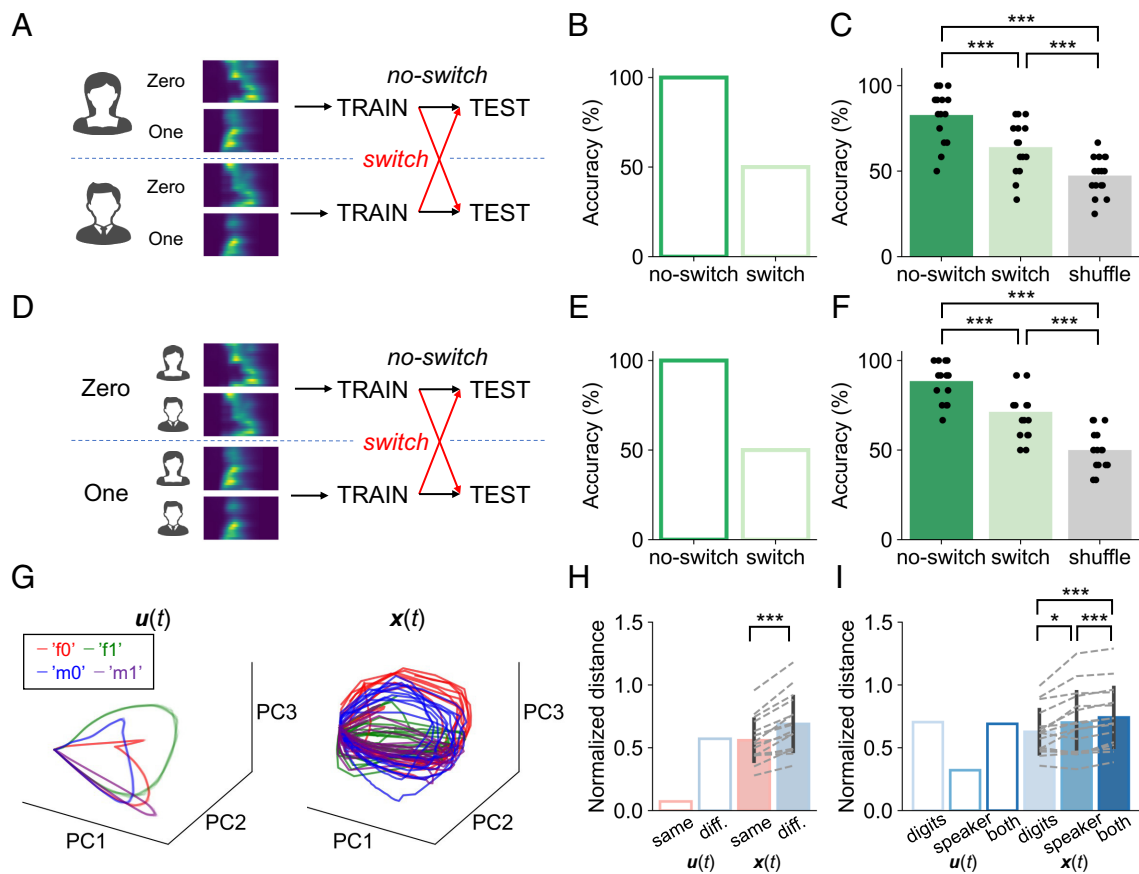


Fig. 5. Generalization ability of mBNN reservoirs. (A) Schematic illustration of the generalized spoken digit classification task. The output weight W is first trained to classify the two digits in one gender, and the accuracy of classifying the digits spoken by a different gender is evaluated during the test phase. (B) The classification accuracy when the photostimulation patterns are directly decoded by a linear decoder. (C) The classification accuracy by the mBNN reservoir. The accuracy is highest when the same inputs are used in the training and testing phases ("no switch"). However, mBNN still retains an accuracy above chance level ("shuffled") even when the gender of the speaker is switched in the training and testing phases ("switch"). *** $P < 0.001$ (two-sided paired t test, $n = 16$, $df = 15$). (D–F) Same as in (A–C) except for the gender classification task. *** $P < 0.001$ (two-sided paired t test, $n = 16$, $df = 15$). (G) Trajectories of the photostimulation pattern $u(t)$ (Left) and reservoir dynamics $x(t)$ (Right) visualized in a three-dimensional principal component subspace. m (male) and f (female) denote the gender, and 0 and 1 denote the digit. The explained variance ratios are 0.94, 0.03, and 0.01 for PC1, PC2, and PC3, respectively, for $u(t)$ (Left) and are 0.73, 0.14, 0.06 for PC1, PC2, and PC3, respectively, for $x(t)$ (Right). (H) Time-averaged distance between the trajectories of input pattern $u(t)$ and reservoir dynamics $x(t)$ for the same (D_{same}) and different input classes (D_{diff}). Bar heights show the mean, and error bars show the SD. *** $P < 0.001$ (two-sided paired t test). (I) Time-averaged distance between the trajectories $u(t)$ and $x(t)$ of different digits (D_{digits}), different speakers (D_{speakers}), and different digits and speakers (D_{both}). * $P < 0.05$; *** $P < 0.001$ (two-sided paired t test, $n = 16$, $df = 15$).

reservoir computing system based on artificial neural networks takes advantage of a sparsely connected random network as the reservoir layer to exploit its high-dimensional dynamics (20, 52). BNNs grown on uniform substrates similarly form random networks, although some metric dependence exists due to the finite length of the axons. However, the connectivity between neurons is usually dense, and spontaneous and evoked activity patterns of BNNs in uniform culture are typically governed by the so-called network bursts, i.e., repetitive firing of neurons that occur coherently across a large population of neurons in a network (53–55). Inducing modular topology in BNNs has been shown to break the excessive coherence in cultured BNNs and increase the variety of intrinsic activity patterns (47, 48, 56), which is probably the underlying mechanism behind the positive correlation between functional modularity and classification accuracy in mBNNs.

We investigated the short-term memory of the mBNN reservoir using a timer task and showed that mBNNs retain a short-term memory of several 100 ms. This value is mostly in agreement with the report by Dranias et al., which showed that the short-term memory in uniformly cultured BNNs in a stimulus-specific memory task was ~ 1 s (57). Kubota et al. reported a memory capacity of 20 ms in uniformly cultured BNNs (58), but the difference of that value

from that in Dranias et al. is most likely caused by the task design. Despite the differences in the reports, characteristic timescales of tens of milliseconds to several seconds, which are comparable to physical reservoirs based on soft materials (28), organic molecules (29), or water in a bucket (59), have been consistently observed in BNNs. This timescale matches that of many real-world signals, such as visual and auditory signals, making BNNs a suitable substrate for physical reservoirs in practical applications.

The mBNN coupled with an external linear decoder was successfully demonstrated to be capable of classifying not only static spatial patterns but also spatiotemporal signals, such as human spoken digits. This contrasts with a previous pioneering study in which a nonlinear classifier (support vector machine) was used to classify spatiotemporal data such as jitter spike train and random music delivered to a uniform BNN (32). While it remains to be investigated whether modularity in BNNs increases performance in other tasks in reservoir computing, such as motor control (33, 60), signal generation (21), and time-series prediction tasks (24), previous computational studies on artificial and spiking neural networks bearing modular structures (34, 36, 38) strongly imply that modularity in mBNN can improve the performance of these tasks as well.

In certain tasks, a higher classification accuracy could be achieved when the output layer was trained to decode directly from the input photostimulation patterns bypassing the mBNN (Fig. 5 *B* and *E* and *SI Appendix*, Fig. S4*B*). The performance of such a decoder, however, was strongly task-dependent. As shown in *SI Appendix*, Fig. S4*B*, the performance suddenly dropped to a chance level when the decoder was tasked to classify signals with identical spatial components. The mBNN reservoir, in contrast, could be trained to distinguish temporally reversed “zero” from the original “zero” (*SI Appendix*, Fig. S4*C*). Furthermore, since the output generated from the photostimulation patterns abruptly decreased to ~ 0 within 1 s (*SI Appendix*, Fig. S7), it would be difficult to train the decoder to generate an output with a delay of more than 1 s. The computational benefit of filtering the signal through BNNs could potentially become evident in such “delayed-output” tasks as well.

Most importantly, our experiments demonstrated that mBNNs act as a generalization filter by reducing the trajectory distance between common classes and increasing the distance between separate classes, enabling classification by readout layers pretrained on different genders or digits. The ability to generalize input signals and encode categorized information is one of the characteristics of information processing in the brain (45), while being a major challenge for machine learning systems (61). Such “categorical learning” was impossible when the input signal was directly decoded without an mBNN. In machine learning, data augmentation techniques such as noise addition, scaling, and rotation have been used to improve the generalization ability by reducing overfitting and expanding the decision boundary of the trained model (62). Although data augmentation has mostly been used for image classification in artificial neural networks, more recently, the technique has been reported to improve the generalization capability in time-series classification (63, 64). We hypothesize that the inherent noise in biological neurons intrinsically performs data augmentation to improve the generalization capability of the BNN reservoir.

Reservoir computing is an elegant framework that enables the establishment of a link between the high-dimensional dynamics of complex systems and their computational functions. By applying the concept of physical reservoirs, we constructively investigated how BNNs transform input signals and how their input responses can be utilized for computing. The mBNN reservoir could pave the way toward providing a platform to investigate structure-function relationships in BNNs, develop brain-like physical reservoirs, and understand the mechanisms of diseases with abnormalities in network modularity and population dynamics, such as autism (65) and epilepsy (66).

Materials and Methods

Fabrication of Micropatterned BNNs. mBNNs were constructed by culturing primary neurons on microfabricated coverslips prepared by microcontact printing. A glass coverslip (C018001, Matsunami) was first sonicated in 100% ethanol and Milli-Q water, dried under N_2 flow, and subsequently treated in air plasma for 60 s (PM-100, Yamato). The surface of the coverslip was then rendered nonpermissive to cell adhesion by coating with a 0.2% solution of poly(2-methacryloyloxyethyl phosphorylcholine-co-3-methacryloxypropyl triethoxysilane) (67) in ethanol for 10 s, drying in an ethanol environment for 20 min, baking in an oven at 70 °C for 4 h, and further drying in a vacuum chamber overnight. The coated coverslip was rinsed in ethanol to remove excess physisorbed molecules and sterilize the sample.

The protein ink used for microcontact printing was a mixture of extracellular matrix gel (E1270, Sigma-Aldrich; 1:100 dilution) and poly-D-lysine (50 $\mu\text{g mL}^{-1}$; P0899, Sigma-Aldrich) and was patterned using a polydimethylsiloxane (PDMS) stamp. The fabrication of the PDMS stamp and stamping procedure have been detailed previously (53, 68). Each micropattern comprised four small squares (200 $\mu\text{m} \times 200 \mu\text{m}$) connected by a thin line of approximately 5 to 10 μm . After

stamping the protein ink, four pieces of PDMS (2 mm \times 2 mm \times 0.5 mm) were placed at the periphery of the coverslip as spacers. After drying the patterned protein ink overnight on a clean bench, the coverslip was finally immersed in neuronal plating medium [minimum essential medium (MEM; 11095-080, Gibco) + 5% fetal bovine serum + 0.6% D-glucose]. In some experiments, PDMS microfluidic films bearing identical micropattern structures were fabricated and attached to a poly-D-lysine coated coverslip, as described previously (48), and were used to replace microcontact printing for neuronal patterning.

Cell Culture and Transfection. All procedures involving animal and gene transfection experiments were approved by the Tohoku University Center for Laboratory Animal Research, Tohoku University (approval number: 2020AmA-001), and Tohoku University Center for Gene Research (approval number: 2019AmLMO-001). Primary neurons were obtained from the cortices of embryonic day 18 pups, plated on the microfabricated coverslip, and cocultured with astrocyte feeder cells in 5 mL of N2 medium [MEM + N2 supplement + ovalbumin (0.5 mg mL^{-1}) + 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)]. At 4 d in vitro (DIV), half of the medium was removed, and a red-shifted channelrhodopsin ChrimsonR (Addgene viral prep #59171-AAV1 or #59171-AAV9) using adeno-associated virus vectors was added at a concentration of 2.0 μL . ChrimsonR transiently increases its conductance for cations upon irradiation of red light (69). Therefore, neurons transfected with ChrimsonR transiently increase its spiking probability upon light irradiation, and the degree of increase positively correlates with the amount of irradiated light, which can be modulated, e.g., by area of photostimulation. At 5 DIV, the medium that was removed at 4 DIV was added back, together with 1 mL Neurobasal medium [Neurobasal (21103-049, Gibco) + 2% B-27 supplement (17504-044, Gibco) + 1% GlutaMAX-I (35050-061, Gibco)]. Half of the medium was replaced with the Neurobasal medium at 8 DIV.

Optogenetic Stimulation. Stimulation of cultured neuronal networks was performed using optogenetics by activating ChrimsonR with patterned light. The Patterned light was created using a digital mirror device (DMD; Mightex Polygon400-G) coupled to a light-emitting diode (LED) (Thorlabs Solis 623C; nominal wavelength, 623 nm) via a liquid light guide. The DMD was mounted on an inverted microscope (Olympus IX-83), and the patterned light was reflected onto the sample stage using a short-pass dichroic mirror with an edge wavelength of 556 nm (Semrock FF556-SDI01). Either one, two, or three networks were selected from one coverslip, and for each network, a session of the spatial pattern classification task, a session of the spoken digit classification task, or two sessions of both tasks were conducted. In some experiments, spontaneous activity was recorded for 6 min prior to recording the stimulation responses.

For the spatial pattern classification and the timer tasks, three distinct positions with identical areas (167 \times 200 μm^2) were selected from the cultured neuronal network and irradiated for 100 ms with 10-s intervals. During the interval, the fluorescence intensity of the neurons decayed back to each baseline. The photostimulation patterns were created and uploaded to Polygon400-G using PolyScan2 software (Mightex), and the switching of the patterns was controlled by externally generated transistor-transistor logic signals. The three patterns were delivered 30 times in random order in each session. In case the neurons formed aggregates and were not uniformly distributed within a single module, the photostimulation area was manually adjusted to be $(2/3)a \times a \mu\text{m}^2$, where a is the diameter of the cell aggregate.

For the spoken digit classification task, photostimulation patterns were designed using a custom MATLAB script. First, the spoken digit data of a female and a male saying “zero” and “one” were selected from the TI-46 dataset and converted to a cochleagram by Lyon cochlear filter using the following the filter setting: decimation factor = 250, quality factor = 2, filter stepping factor = 0.25, and number of frequency channels = 17. The same four waveforms were used in the training and test phases. To convert the cochleagram to photostimulation patterns, each of the frequency channels in the cochleagram was first mapped one-to-one to individual neurons expressing ChrimsonR, which were selected randomly from two neighboring modules. The spectrum intensity of the cochleagram was then normalized by the maximum value in the entire cochleagram and converted to circular irradiation areas such that the value of the normalized cochleagram $v \in [0, 1]$ corresponds to a diameter of $25 \times v \mu\text{m}$. Finally, black-and-white bitmap files containing the photostimulation patterns were generated and imported into PolyScan2 software.

The photostimulation patterns corresponding to “zero” and “one” were alternately irradiated every 15 s for a total of 20 times in each session.

Calcium Imaging. At 10 to 12 DIV, the coverslips with mBNN were rinsed in HEPES-buffered saline (HBS) containing 128 mM NaCl, 4 mM KCl, 1 mM CaCl₂, 1 mM MgCl₂, 10 mM D-glucose, 10 mM HEPES, and 45 mM sucrose, and subsequently loaded with 2 μM fluorescent calcium indicator Cal-520 AM (AAT Bioquest) in HBS for 30 min in an incubator. The coverslip was then transferred to a glass-bottom dish (3960-035, Iwaki) filled with fresh HBS. The dynamics of the calcium indicator fluorescence in response to optogenetic stimulation were imaged using an inverted microscope (IX83, Olympus) equipped with a 20× objective lens (numerical aperture, 0.75), an LED light source (Lambda HPX, Sutter Instrument), a scientific complementary metal-oxide semiconductor camera (Zyla 4.2P, Andor), and an incubation chamber (Tokai Hit). All the recordings were performed at 37 °C. Up to three mBNNs were selected from a coverslip, and fluorescence imaging was performed at 20 frames/s using the Solis software (Andor).

Reservoir Computing. To use the dynamics of an mBNN as the reservoir state, N (=60) neurons (15 neurons per module) were manually selected from a single network, and their somas were defined as regions of interest (ROIs). A fraction of the ROIs was selected from the modules receiving optogenetic stimulation. The RFU $\Delta F/F$ of cell i at time t was calculated as $[F_i(t) - F_0]/F_0$, where $F_i(t)$ is the mean intensity within ROI i at time t and F_0 is the baseline fluorescence of the ROI. $\Delta F/F$ was then used to construct the network dynamics $\mathbf{x}(t)$ ($= [x_i(t)]$) as $x_i(t) = [F_i(t) - F_0]/F_0$. For the human spoken digit classification task, in order to compensate for the fluctuation in the frame rate in the DMD, we posed an assumption that the rise time of the neuron-averaged RFU was consistent across different repetitions and classes, and regarded the stimulus onset to be at 0.5 s prior to the peak of the neuron-averaged RFU.

The reservoir computing model consisted of an input layer, a reservoir layer, and a readout layer. Input $\mathbf{u}(t)$ was delivered optogenetically, as described above. The reservoir layer was an mBNN that generated an N -dimensional signal $\mathbf{x}(t)$, which was determined by the intrinsically generated spontaneous activity and by its response to $\mathbf{u}(t)$. The readout layer was an N_y -dimensional linear decoder, with $N_y = 3, 1$, and 2 for the spatial pattern classification, timer, and spoken digit classification tasks, respectively. The readout state at time t , $\mathbf{y}(t)$, is calculated as:

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t), \quad [1]$$

where \mathbf{W} denotes the output weight matrix.

In the spatial pattern (spoken digit) classification tasks, datasets from the first 20 (14) trials were used for training and those from the subsequent 10 (6) trials were used for testing. For the training phase, \mathbf{W} was obtained using ridge regression as:

$$\mathbf{W} = \hat{\mathbf{Y}}\mathbf{X}^T(\mathbf{X}\mathbf{X}^T + \lambda\mathbf{I})^{-1}, \quad [2]$$

where \mathbf{X} is the reservoir state matrix, $\hat{\mathbf{Y}}$ is the target signal matrix, λ (=1) is the regularization coefficient, and \mathbf{I} is an identity matrix. \mathbf{X} is derived by horizontally concatenating $\mathbf{x}(t)$ from the onset of the first trial to the end of the last trial during the training phase. $\hat{\mathbf{Y}}$ is derived similarly by concatenating $\hat{\mathbf{y}}(t)$, which is an N_y -dimensional target function vector with an element corresponding to the correct label set to one for a duration of 2.5 s from each stimulus onset and zero otherwise. For the spatial pattern classification task, the entire time steps during the training phase are used to generate \mathbf{X} and $\hat{\mathbf{Y}}$. For the spoken digit classification task, the first 10 s from the stimulus onset is extracted from each trial and are used to generate the matrices.

During the test phase of the two classification tasks, the output vector $\mathbf{y}(t)$ is obtained using Eq. 1 with a fixed weight matrix \mathbf{W} derived by Eq. 2. Subsequently, for each mBNN, the fraction of correct estimates (classification accuracy) is calculated by the number of test trials that gave a correct answer divided by the total number of test trials (10 and 6 for the spatial pattern and spoken digit classification tasks, respectively). The estimated answer was obtained from $\mathbf{y}(t)$ as $\arg\max_{i \in N_y} \left[\sum_{t=t_k}^{t_k + \Delta t} y_i(t) \right]$, where i is an element in the output layer, t_k is the onset

of the k th input, and Δt is the duration of the target (set to 2.5 s). Accuracy was evaluated as the fraction of correct estimates.

In the timer task, the network response $\mathbf{x}(t)$ was regressed to produce a delayed target signal using ridge regression. More precisely, the target output is set equal to a square wave identical to the input signal, whose onset was delayed by τ s

against the onset of stimulation, and an optimized weight matrix (\mathbf{W}_τ) is derived using Eq. 2 for each τ in each of the 30 trials. For each \mathbf{W}_τ , a delayed output $y_\tau(t)$ is then obtained as $y_\tau(t) = \mathbf{W}_\tau \mathbf{x}(t)$. The short-term memory of the neuronal network was evaluated using the coefficient of determination $r^2(\tau)$ as (50):

$$r^2(\tau) = \frac{\text{Cov}^2[u(t - \tau), y_\tau(t)]}{\text{Var}[u(t)] \cdot \text{Var}[y_\tau(t)]}, \quad [3]$$

where Cov is the covariance, Var the variance, and $u(t)$ the input. The inputs were separated by an interval of 10 s to allow the mBNN return to its baseline state. Since the output layer is memoryless, the task evaluates the temporal property of the reservoir layer.

Trajectory Analysis. To quantify the similarity/dissimilarity between the trajectories, the normalized distance between two trajectories $\mathbf{p}(t)$ and $\mathbf{q}(t)$, D_{pq} , was calculated as follows:

$$D_{pq} = \langle d_{pq}(t) \rangle_t = \frac{1}{\sqrt{L}} \langle \|\mathbf{p}_z(t) - \mathbf{q}_z(t)\|_2 \rangle_t, \quad [4]$$

where $\mathbf{p}_z(t)$ is $\mathbf{p}(t)$ standardized over all time points and all nodes, L is the number of nodes, $d_{pq}(t)$ is the normalized distance between trajectories \mathbf{p} and \mathbf{q} at time t , $\|\cdot\|_2$ and $\langle \cdot \rangle_t$ denote the L2 norm and temporal average, respectively. The temporal average was calculated over 5 s of a trajectory, starting from the stimulus onset. The mean distance of the trajectories in response to the same (different) input classes was defined as D_{same} (D_{diff}) and evaluated as follows:

$$D_{\text{same}} = \langle D_{pq} \rangle, (c_p = c_q),$$

$$D_{\text{diff}} = \langle D_{pq} \rangle, (c_p \neq c_q), \quad [5]$$

where c_i is the input class of trajectory i ($p \neq q$), and $\langle \cdot \rangle$ is the mean across the corresponding trajectory pairs.

Network Analysis. For the analysis of functional connectivity within mBNNs, the firing rate of neuron i , f_i , was inferred from the relative fluorescence intensity $\Delta F/F$ using the CASCADE algorithm with a convolutional neural network trained with the Global_EXC dataset (sampling rate = 20 Hz, smoothing window = 100 ms) (70). The Pearson correlation coefficient between neurons i and j , r_{ij} , was calculated as follows:

$$r_{ij} = \frac{\text{Cov}[f_i(t), f_j(t)]}{\sqrt{\text{Var}[f_i(t)] \cdot \text{Var}[f_j(t)]}}. \quad [6]$$

The covariance and variance were calculated using the data from spontaneous activity recordings, eliminating the first and last 30 s of each recording. Following Sporns et al. (49), the functional modularity Q of the correlation matrix was then quantified using Newman's method (71) as follows:

$$Q = \frac{1}{2M} \sum_{ij} \left(r_{ij} - \frac{k_i k_j}{2M} \right) \delta(m_i, m_j), \quad [7]$$

where M is the sum of all weighted edges ($= \frac{1}{2} \sum_{ij} r_{ij}$), k_i is the sum of the weighted edges attached to neuron i ($= \sum_j r_{ij}$), δ is the Kronecker delta, and m_i is the module containing neuron i .

Statistical Statement. All values in the analyzed data are presented as mean \pm SD or as means. Sample size (n) and degrees of freedom (df) are also presented at each section in the text and figure captions. Statistical significance of the mean values between two groups were compared using two-sided paired t tests, while correlations between two variables were tested using the two-sided Pearson correlation test.

Data, Materials, and Software Availability. Datasets generated and analyzed during the current study are archived in Zenodo at DOI: [10.5281/zenodo.7792577](https://doi.org/10.5281/zenodo.7792577) (72).

ACKNOWLEDGMENTS. We thank Kei Wakimura and Taiki Takemuro for setting up the fluorescence microscope. We acknowledge the Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT) Grant-in-Aid for Transformative Research

Areas (B) "Multicellular Neurobiocomputing" (21H05163 and 21H05164), the Japan Society for the Promotion of Science (JSPS) KAKENHI (18H03325, 19H00846, 20H02194, 20K20550, 21J21766, 22H03657, 22K19821, and 22KK0177), the Japan Science and Technology Agency (JST) PRESTO (JMPJPR18MB), JST CREST (JPMJCR19K3), and Tohoku University Research Institute of Electrical Communication (RIEC) Cooperative Research Project Program for financial support. Cartoons in Figs. 1 A and B and 4A were created with [BioRender.com](https://www.biorender.com).

1. D. V. Buonomano, W. Maass, State-dependent computations: Spatiotemporal processing in cortical networks. *Nat. Rev. Neurosci.* **10**, 113–125 (2009).
2. A. Luczak, B. L. McNaughton, K. D. Harris, Packet-based communication in the cortex. *Nat. Rev. Neurosci.* **16**, 745–755 (2015).
3. S. Saxeena, J. P. Cunningham, Towards the neural population doctrine. *Curr. Opin. Neurobiol.* **55**, 103–111 (2019).
4. S. Vyas, M. D. Golub, D. Sussillo, K. V. Shenoy, Computation through neural population dynamics. *Annu. Rev. Neurosci.* **43**, 249–275 (2020).
5. L. Duncker, M. Sahani, Dynamics on the manifold: Identifying computational dynamical activity from neural population recordings. *Curr. Opin. Neurobiol.* **70**, 163–170 (2021).
6. W. Singer, Recurrent dynamics in the cerebral cortex: Integration of sensory evidence with stored knowledge. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2101043118. (2021).
7. K. Diba, H. A. Lester, C. Koch, Intrinsic noise in cultured hippocampal neurons: Experiment and modeling. *J. Neurosci.* **24**, 9723–9733 (2004).
8. R. J. Sayer, M. J. Friedlander, S. J. Redman, The time course and amplitude of EPSPs evoked at synapses between pairs of CA3/CA1 neurons in the hippocampal slice. *J. Neurosci.* **10**, 826–836 (1990).
9. J.-E. K. Miller, I. Ayzenshtat, L. Carrillo-Reid, R. Yuste, Visual stimuli recruit intrinsically generated cortical ensembles. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E4053–E4061 (2014).
10. J. S. Montijn, G. T. Meijer, C. S. Lansink, C. M. A. Pennartz, Population-level neural codes are robust to single-neuron variability from a multidimensional coding perspective. *Cell Rep.* **16**, 2486–2498 (2016).
11. L. Carrillo-Reid, J.-E. K. Miller, J. P. Hamm, J. Jackson, R. Yuste, Endogenous sequential cortical activity evoked by visual stimuli. *J. Neurosci.* **35**, 8813–8828 (2015).
12. J. S. Montijn, U. Olcese, C. M. A. Pennartz, Visual stimulus detection correlates with the consistency of temporal sequences within stereotyped events of V1 neuronal population activity. *J. Neurosci.* **36**, 8624–8640 (2016).
13. A. Luczak, P. Barthó, K. D. Harris, Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron* **62**, 413–425 (2009).
14. M. M. Churchland *et al.*, Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).
15. J. A. Michaels, S. Schaffelhofer, A. Agudelo-Toro, H. Scherberger, A goal-driven modular neural network predicts parietofrontal neural dynamics during grasping. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 32124–32135 (2020).
16. C. D. Harvey, P. Coen, D. W. Tank, Choice-specific sequences in parietal cortex during a virtual navigation decision task. *Nature* **484**, 62–68 (2012).
17. D. V. Buonomano, M. M. Merzenich, Temporal information transformed into spatial code by a neural network with realistic properties. *Science* **267**, 1028–1030 (1995).
18. H. Jaeger, *The "Echo State" Approach to Analysing and Training Recurrent Neural Networks* (German National Research Center for Information Technology, 2001).
19. W. Maass, T. Natschlager, H. Markram, Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Comput.* **14**, 2531–2560 (2002).
20. M. Lukoševičius, H. Jaeger, Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **3**, 127–149 (2009).
21. D. Sussillo, L. F. Abbott, Generating coherent patterns of activity from chaotic neural networks. *Neuron* **63**, 544–557 (2009).
22. G. Tanaka *et al.*, Recent advances in physical reservoir computing: A review. *Neural Netw.* **115**, 100–123 (2019).
23. K. Nakajima, Physical reservoir computing—An introductory perspective. *Jpn. J. Appl. Phys.* **59**, 065051 (2020).
24. D. Brunner, M. C. Soriano, C. R. Mirasso, I. Fischer, Parallel photonic information processing at gigabyte per second data rates using transient states. *Nat. Commun.* **4**, 1364 (2013).
25. S. Sunada, A. Uchida, Photonic reservoir computing based on nonlinear wave dynamics at microscale. *Sci. Rep.* **9**, 19078 (2019).
26. J. Torreon *et al.*, *Nature* **547**, 428–431 (2017).
27. J. Grollier *et al.*, *Nat. Electron.* **3**, 360–370 (2020).
28. K. Nakajima, T. Li, H. Hauser, T. Li, R. Pfeifer, Exploiting short-term memory in soft body dynamics as a computational resource. *J. R. Soc. Interface* **11**, 20140437 (2014).
29. Y. Usami *et al.*, In-materio reservoir computing in a sulfonated polyaniline network. *Adv. Mater.* **33**, 2102688 (2021).
30. M. E. Ruaro, P. Bonifazi, V. Torre, Toward the neurocomputer: Image processing and pattern recognition with neuronal cultures. *IEEE Trans. Biomed. Eng.* **52**, 371–383 (2005).
31. K. P. Dockendorf, I. Park, P. He, J. C. Principe, T. B. DeMarse, Liquid state machines and cultured cortical networks: The separation property. *BioSystems* **95**, 90–97 (2009).
32. H. Ju, M. R. Dranias, G. Banumurthy, A. M. J. VanDongen, Spatiotemporal memory is an intrinsic property of networks of dissociated cortical neurons. *J. Neurosci.* **35**, 4040–4051 (2015).
33. Y. Yada, S. Yasuda, H. Takahashi, Physical reservoir computing with FORCE learning in a living neuronal culture. *Appl. Phys. Lett.* **119**, 173701 (2021).
34. N. Rodriguez, E. Izquierdo, Y. Y. Ahn, Optimal modularity and memory capacity of neural reservoirs. *Netw. Neurosci.* **3**, 551–566 (2019).
35. S. Moriya, H. Yamamoto, A. Hirano-Iwata, S. Kubota, S. Sato, Quantitative analysis of dynamical complexity in cultured neuronal network models for reservoir computing applications. In *2019 International Joint Conference on Neural Networks (IJCNN) N-20275* (IEEE, 2019).
36. Y. Kawai, J. Park, M. Asada, A small-world topology enhances the echo state property and signal propagation in reservoir computing. *Neural Netw.* **112**, 15–23 (2019).

Author affiliations: ^aResearch Institute of Electrical Communication, Tohoku University, Sendai 980-8577, Japan; ^bGraduate School of Biomedical Engineering, Tohoku University, Sendai 980-8579, Japan; ^cSchool of Engineering, Tohoku University, Sendai 980-8579, Japan; ^dGraduate School of Systems Information Science, Future University Hakodate, Hakodate 041-8655, Japan; ^eInstitute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan; ^fGraduate School of Pharmaceutical Sciences, Tohoku University, Sendai 980-8578, Japan; and ^gWorld Premier International Research Center Initiative—Advanced Institute for Materials Research, Tohoku University, Sendai 980-8577, Japan

37. Z. Deng, Y. Zhang, Collective behavior of a small-world recurrent neural system with scale-free distribution. *IEEE Trans. Neural Netw.* **18**, 1364–1375 (2007).
38. L. E. Suarez, B. A. Richards, G. Lajoie, B. Misic, Learning function from structure in neuromorphic networks. *Nat. Mach. Intell.* **3**, 771–786 (2021).
39. M. A. Bertolo, B. T. T. Yeo, M. D'Esposito, The modular and integrative functional architecture of the human brain. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E6798–E6807 (2015).
40. T. Murakami, T. Matsui, M. Uemura, K. Ohki, Modular strategy for development of the hierarchical visual network in mice. *Nature* **608**, 578–585 (2022).
41. C. Zhou, L. Zemanová, G. Zamora, C. C. Hilgetag, J. Kurths, Hierarchical organization unveiled by functional connectivity in complex brain networks. *Phys. Rev. Lett.* **97**, 238103 (2006).
42. W.-C. A. Lee *et al.*, Anatomy and function of an excitatory network in the visual cortex. *Nature* **532**, 370–374 (2016).
43. D. Meunier, R. Lambiotte, E. T. Bullmore, Modular and hierarchically modular organization of brain networks. *Front. Neurosci.* **4**, 200 (2010).
44. M. P. van den Heuvel, E. T. Bullmore, O. Sporns, Comparative connectomics. *Trends Cogn. Sci.* **20**, 345–361 (2016).
45. S. Reinert, M. Hubener, T. Bonhoeffer, P. M. Goltstein, Mouse prefrontal cortex represents learned rules for categorization. *Nature* **593**, 411–417 (2021).
46. A. Uchida, R. McAllister, R. Roy, Consistency of nonlinear system response to complex drive signals. *Phys. Rev. Lett.* **93**, 244102 (2004).
47. H. Yamamoto *et al.*, Impact of modular organization on dynamical richness in cortical networks. *Sci. Adv.* **4**, eaau4914 (2018).
48. T. Takemuro, H. Yamamoto, S. Sato, A. Hirano-Iwata, Polydimethylsiloxane microfluidic films for in vitro engineering of small-scale neuronal networks. *Jpn. J. Appl. Phys.* **59**, 117001 (2020).
49. O. Sporns, R. F. Betzel, Modular brain networks. *Annu. Rev. Psychol.* **67**, 613–640 (2016).
50. H. Jaeger, Short-term memory in echo state networks. *GMD Report* **152** (2002).
51. P. Mujal *et al.*, Opportunities in quantum reservoir computing and extreme learning machines. *Adv. Quantum Technol.* **4**, 2100027 (2021).
52. Y. Yonemura, Y. Katori, Network model of predictive coding based on reservoir computing for multimodal processing of visual and auditory signals. *NOLTA, IEICE* **12**, 143–156 (2021).
53. H. Yamamoto *et al.*, Size-dependent regulation of synchronized activity in living neuronal networks. *Phys. Rev. E* **94**, 012407 (2016).
54. V. Pasquale, S. Martinoia, M. Chiappalone, Stimulation triggers endogenous activity patterns in cultured cortical networks. *Sci. Rep.* **7**, 9080 (2017).
55. M. Dazza, S. Metens, P. Monceau, S. Bottani, A novel methodology to describe neuronal networks activity reveals spatiotemporal recruitment dynamics of synchronous bursting states. *J. Comput. Neurosci.* **49**, 375–394 (2021).
56. M. U. Park *et al.*, Collective dynamics of neuronal activities in various modular networks. *Lab Chip* **21**, 951–961 (2021).
57. M. R. Dranias, H. Ju, E. Rajaram, A. M. VanDongen, Short-term memory in networks of dissociated cortical neurons. *J. Neurosci.* **33**, 1940–1953 (2013).
58. T. Kubota, K. Nakajima, H. Takahashi, "Echo state property of neuronal cell cultures" in *Lecture Notes in Computer Science 11731* (Springer, Berlin, 2019), pp. 137–148.
59. C. Fernando, S. Sojakka, "Pattern recognition in a bucket" in *Lecture Notes in Computer Science 2801*, (Springer, Berlin, 2003), pp. 588–597.
60. Y. Yoshino, Y. Katori, Short-term memory ability of reservoir-based temporal difference learning model. *NOLTA, IEICE* **13**, 203–208 (2022).
61. F. H. Sinz, Z. Pitkow, J. Reimer, M. Bethge, A. S. Tolia, Engineering a less artificial intelligence. *Neuron* **103**, 967–979 (2019).
62. C. Shorten, T. M. Khoshgoftaar, A survey on image data augmentation for deep learning. *J. Big Data* **6**, 60 (2019).
63. S. Demir, K. Mincev, K. Kok, N. G. Paterakis, Data augmentation for time series regression: Applying transformations, autoencoders and adversarial networks to electricity price forecasting. *Appl. Energy* **304**, 117695 (2021).
64. B. K. Iwana, S. Uchida, An empirical survey of data augmentation for time series classification with neural networks. *PLoS ONE* **16**, e0254841. (2021).
65. D. G. Amaral, C. M. Schumann, C. W. Nordahl, Neuroanatomy of autism. *Trends Neurosci.* **31**, 137–145 (2008).
66. C. J. Stam, Modern network science of neurological disorders. *Nat. Rev. Neurosci.* **15**, 683–695 (2014).
67. Y. Xu, M. Takai, T. Konno, K. Ishihara, Microfluidic flow control on charged phospholipid/polymer interface. *Lab Chip* **7**, 199–206 (2007).
68. H. Yamamoto *et al.*, Unidirectional signal propagation in primary neurons micropatterned at a single-cell resolution. *Appl. Phys. Lett.* **109**, 043703 (2016).
69. N. C. Klapoetke *et al.*, Independent optical excitation of distinct neural populations. *Nat. Meth.* **11**, 338–346 (2014).
70. P. Ruppert *et al.*, A database and deep learning toolbox for noise-optimized, generalized spike inference from calcium imaging. *Nat. Neurosci.* **24**, 1324–1337 (2021).
71. M. E. J. Newman, Analysis of weighted networks. *Phys. Rev. E* **70**, 056131 (2004).
72. T. Sumi *et al.*, Dataset for: Biological neurons act as generalization filters in reservoir computing. *Zenodo*. <https://zenodo.org/record/7792577>. Deposited 8 May 2023.