Neurocomputing 671 (2026) 132705

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Temporal-filter enhanced prediction of whole-brain neural activity using the physiology-aligned latent variable model

Chong Li [a,b,c,d,e], Zhongyu Chen [a,b,c,d,e], Jianfeng Feng [c,d], Xiangyang Xue [a,c,d], Yuguo Yu [b,c,d,e,*]

[a] College of Computer Science and Artificial Intelligence, Fudan University, China
[b] Research Institute of Intelligent and Complex Systems, China
[c] MOE Frontiers Center for Brain Science, and State Key Laboratory of Brain Function and Disorders, Fudan University, China
[d] Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai 200433, China
[e] Shanghai Artificial Intelligence Laboratory, Shanghai 200232, China

ARTICLE INFO

Keywords:
Deep neural networks
Temporal-filter
Inference network
Self-focused attention
Physiology-aligned latent variable algorithm
Connectomics
Generative model
Independent coding
C. elegans

ABSTRACT

Understanding the neural mechanisms behind intelligent behaviors in nematodes requires a comprehensive framework that integrates connectome data with recordings of all neuronal activities. However, obtaining full membrane potentials for all labeled neurons is challenging, and even in C. elegans, which has only 302 neurons, current calcium imaging techniques typically capture just over half of these labeled neurons in vivo. Existing advances primarily focus on using spatial correlation across the whole brain to causally predict unknown neuronal activities by using measured activities of half-numbered neurons, but they struggle to achieve high prediction accuracy. By introducing a quasi-independent temporal coding property of populational neurons in living brains, e.g., C. elegans, we establish the Temporal-Filter and Physiology-Aligned Latent Variable Model (TF-PALVM), a new algorithmic leap that synergizes the spatial connectome structure with temporal coding functions of individual neurons to infer the electrical activities of the entire neural ensemble. This model employs an autoencoder network with temporal kernels, rather than relying on spatial correlations, refined to reflect individual neuronal temporal coding functions and predict their future temporal activities, while embedding experimentally derived synaptic weights into a biologically coherent framework. When tested, our model demonstrates unprecedented reconstruction accuracy, surpassing existing models by approximately 75 % in the worm holdout evaluation and 51 % in neuron holdout performance. Moreover, it precisely predicts synaptic polarities, with 75 % of them to be excitatory, matching experimental excitatory synapse data. It is also able to identify the top neuron pairs with the most influence on behavioral correlations, consistent with previous experimental research. Our TF-PALVM stands as a transformative tool for neuroscientific exploration, capable of predicting missing neuronal activity with high fidelity. The success of the model confirms that the temporal response history of individual neurons contains more valuable information than the population network in predicting their present and future responses to sensory inputs. It offers a scalable approach to potentially unravel the complexities of larger, more intricate brains.

1. Introduction

Caenorhabditis elegans (C. elegans), renowned for its simple anatomy, genetics, and a compact nervous system comprising merely 302 neurons in the hermaphrodite, is widely employed as a representative model for studying the underpinnings of biological intelligence. Its unique attributes, including a short life cycle, easily accessible mutants, diminutive size, and transparent body, render it an exemplary organism for the development of intricate, brain-inspired artificial intelligence models. A pivotal step towards unraveling the neural control mechanisms underlying intelligent sensory and motor behaviors in C. elegans involves the development of a comprehensive computational model. This model should faithfully replicate the real-world neuronal activities and behaviors inherent in this organism. However, obtaining invasive electronic recordings of complete voltage signals across all neurons in a living nematode presents a significant challenge. Conversely, non-

* Corresponding author at: Research Institute of Intelligent and Complex Systems, China.
E-mail address: yuyuguo@fudan.edu.cn (Y. Yu).

https://doi.org/10.1016/j.neucom.2026.132705
Received 21 June 2024; Received in revised form 10 December 2025; Accepted 14 January 2026
Available online 16 January 2026

invasive calcium fluorescence imaging, widely accepted as a surrogate for voltage measurement, has been instrumental in tracking neural activity. It has enabled simultaneous recording from multiple labeled neurons [1–3]. However, there is still no in-vivo recorded dataset of fully 302 labeled neurons available till now, while the most number of available datasets published is around 170 labeled neurons [3].

To make the best use of these ~170 recorded data, Mi et al. [4] introduce the Connectome-Constrained Latent Variable Model (CCLVM). This pioneering framework employs a variational autoencoder, emphasizing the use of the global spatial 300×300 matrix of brain connectivity in C. elegans. Meanwhile, they also construct a biophysical network model with previously published connectome data [5]. Then they train the CCLVM framework with available calcium activities of ~170 recorded labeled neurons of 21 worms. The model integrates voltage distribution inference, fluorescence reconstruction, together with the generative activities of the connectome-constrained biophysical model. It demonstrates the ability to replicate neuronal voltage and calcium signals of ~170 recorded neurons with accuracy ranging between 0.3 and 0.5, while also predicting neuronal activities of all the 130 unobserved neurons [4].

Although the pioneering CCLVM algorithm demonstrates great potential in reconstructing the entire set of brain neuronal activities, its limited prediction accuracy suggests an intrinsic bottleneck. Upon examining their algorithm, it was noted that the inference network's autoencoder relies heavily on the spatial correlations among all neurons in C. elegans, employing recorded calcium activities to predict missing neuronal activities. However, broader neuroscience studies have suggested that neurons in biological brains function largely as independent encoders rather than as part of a population encoder [6,7]. This implies that the historical temporal responses of individual neurons to signals, rather than spatial correlations across populations, have a greater impact on predicting the present spiking response of each individual neuron. This potentially mitigates the limitations imposed by the small number of neurons available and indicates the independence of neurons [8–10]. Additionally, to further improve the comparability of model predictions with realistic neuronal recording data from experiments, it is advisable to use reasonable physiological parameters. These include ionic channel conductance with standard units, resting and depolarized membrane potential ranges, synaptic polarity, and reasonable reverse potential parameters, among others [11–13]. Therefore, in this study, to achieve state-of-the-art reconstruction and prediction accuracy, we first establish a new inference framework by combining the temporal filter properties of individual neurons with the spatial connectome structure of C. elegans, called the Temporal-Filter Inference Network (TFIN). Furthermore, we incorporate realistic voltage and conductance units as well as prior physiological knowledge into the worm neural network, to make the model more biological plausibility and improving congruence with the actual nematode. This creates a brand-new refined model, i.e., Temporal-Filter and Physiology-Aligned Latent Variable Model (TF-PALVM). Specifically, on the one hand, we integrate temporal convolutional layers with designated temporal kernels into the inference network, allowing an independent autoencoder to exclusively focuses on the temporal dimension of neural activity, and on the other hand, we incorporate voltage units and prior knowledge into the worm neural network, enhancing biological plausibility and improving congruence with the actual nematode. A summary comparison between TF-PALVM and representative methods is provided in Table 1.

To evaluate model performance, we employ the NeuroPAL dataset [3] for both worm holdout evaluation, aiming to indicate precision of the conversion from fluorescence to voltage, and neuron holdout evaluation, focused on gauging the accuracy of predicting unobserved neuronal activity, and compare the results with experimental data and previous studies [4]. Our results show a significant increase in accuracy for both reconstruction and prediction when incorporating the temporal functions of individual neurons and physiological parameters into both the inference and generative networks. In the worm holdout evaluation, TF-PALVM demonstrates an impressive increase of ~75 % in correlation, signifying its superior ability to capture temporal patterns for independent encoding. Similarly, in the neuron holdout evaluation, TF-PALVM outperforms CCLVM by ~51 % in correlation, underscoring its effectiveness in accurately predicting neuronal activity. Furthermore, leveraging the most comprehensive connectome available [14] and partially measured polarity data [11], we employ TF-PALVM to predict the polarity of each connection in C. elegans, achieving an accurate excitatory synapse ratio of ~75 %, closely aligning with the proportion (~77 %) obtained from biological experimental data [11], while CCLVM's prediction deviates (~49 %). Additionally, we conduct ablation studies to present a neuron sorting analysis and gain insights into the population as well as the independent coding properties of TF-PALVM.

This study introduces several innovative contributions to the field, which are summarized as follows:

1. Our TF-PALVM predicts the electrical activity of all neurons using partially measured calcium fluorescence data with significantly higher accuracy than the pioneer work [4]. We have fundamentally reimagined the inference network architecture by harnessing the independent nature of neuronal encoding [8–10]. By integrating temporal-filter convolutional layers, we achieve a remarkable ~75 % improvement compared to the CCLVM [4] in accurately translating calcium fluorescence data into electrical activity data. This substantial advancement has been validated through our worm holdout evaluation. Further, by incorporating voltage units and prior knowledge based on previous studies [11,13], we have achieved an enhanced prediction accuracy of ~51 % over the CCLVM for the electrical activity of unobserved neurons, as corroborated by our neuron holdout evaluation.
2. We have developed a connectome-constrained neural network that closely emulates the biophysical attributes of real nematodes. The exceptional precision achieved by our model underscores its potential to surpass current methodologies in generating high-quality whole-brain electrical activity data from calcium fluorescence data collected from a select group of neurons. Our model's unique strength lies in its prediction of a chemical synapse excitatory ratio of ~75 %, which closely aligns with the experimentally determined value of ~77 % [11]. In contrast, the CCLVM's prediction (~49 %) deviates considerably from the observed experimental data.
3. We conduct ablation study and neuron sorting analysis to identify the top 5 pairs of neurons with the most impact in predicting other neurons' activity (RMDV, AVA, RME, AVB, ASE), crawling distance (RMED, URYD, RIC, RMD, RMG), and chemotaxis index (ADF, AIA, RMDD, VA01, ASH).

Overall, this study represents a pioneering and potent approach that substantially enhances the precision and biological plausibility of neuronal activity modeling in nematodes.

Table 1
Summary comparison between TF-PALVM and representative models for whole-brain neural activity prediction in C. elegans.

Method	Connectome	Neuron-wise Temporal Filter	Synaptic Polarity	Physiological Units	Prediction Accuracy
LVM[6]	×	×	×	×	low
CC-LVM[6]	✔	×	×	×	medium
TF-PALVM(Ours)	✔	✔	✔	✔	high

1.1. Related work

Neural information encoding has long been studied from both independent and population-based perspectives. Empirical studies suggest that neurons often encode information independently rather than solely through population coding. For example, retinal ganglion cells act largely as independent encoders [7], and sparse, hierarchical coding has been observed in the C. elegans sensory system [10]. Temporal coding, where information is carried by historical activity or timing precision, is also a fundamental principle in neural computation [6]. These findings provide strong biological motivation for modeling neuron-specific temporal dynamics, which directly inspires the use of independent temporal filters in our framework.

With the availability of the full C. elegans connectome [5], several computational frameworks have been developed to incorporate structural connectivity into neural activity modeling. The CCLVM [4] pioneered the integration of connectome data [5] into a variational autoencoder to predict unobserved neuronal activity. While effective, CCLVM relied mainly on spatial correlations and achieved limited accuracy. Subsequent models of the C. elegans nervous system [4] improved realism but still lacked explicit temporal modeling and polarity constraints, limiting their ability to capture neural dynamics.

Recent advances in both neuroscience and machine learning have emphasized the importance of incorporating physiological priors into data-driven models. Studies have shown that incorporating physiological priors such as ion channel parameters, membrane potentials, and synaptic polarities enhances model plausibility [11–13]. In parallel, latent variable models and temporal convolutional networks have been widely used in machine learning for sequential data [4]. TF-PALVM uniquely combines these directions by embedding temporal filters into the inference network and constraining the generative model with updated connectome and polarity data, enabling both improved accuracy and biological interpretability.

2. Materials and methods

2.1. Temporal-filter and physiology-aligned latent variable model

Our Temporal-Filter and Physiology-Aligned Latent Variable Model (TF-PALVM) expands on the variational autoencoder framework introduced in CCLVM [4], as depicted in Fig. 1. Unlike CCLVM, which relies on the spatial correlation property from all neurons to predict each neuron, TF-PALVM incorporates the temporal-filter module in the inference network to independently extract historical information for each neuron. Furthermore, TF-PALVM integrates prior knowledge to impose more robust constraints on biophysical properties of the worm neural network with units, thereby enhancing biological plausibility and ensuring biological authenticity. As a result, this framework is capable to accurately generate the membrane potential signals for all neurons based on partially recorded calcium fluorescence data. This model comprises three main components: the inference network, the reconstruction model, and the worm neural network, elucidated as follows.

2.1.1. Inference network

The inference network $Q(\mathbf{v}|\mathbf{f},\mathbf{o})$ produces the estimated posterior distribution of the latent variable $\mathbf{v}$, given the chemosensory stimulus $\mathbf{o}$

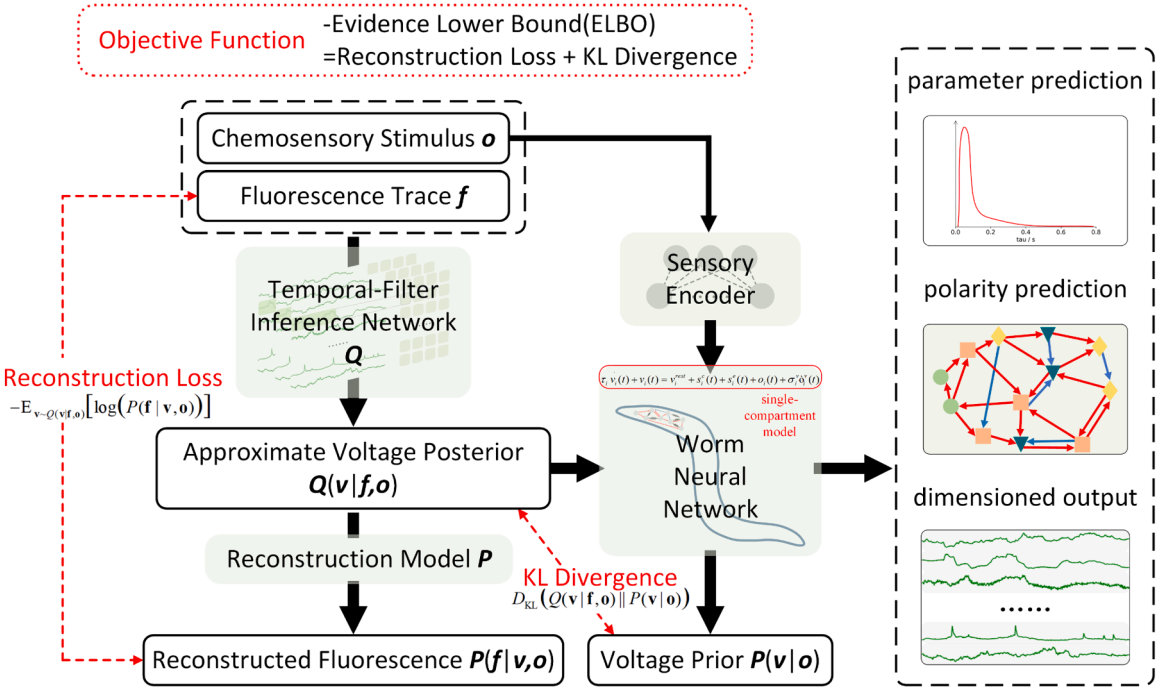


Fig. 1. Framework of the Temporal-Filter and Physiology-Aligned Latent Variable Model. This framework comprises three key components: an inference network $Q(\mathbf{v}|\mathbf{f},\mathbf{o})$, a reconstruction model $P(\mathbf{f}|\mathbf{v},\mathbf{o})$ and a worm neural network $P(\mathbf{v}|\mathbf{o})$. In contrast to CCLVM [4], significant improvements have been made to both the inference network, which now more accurately converts fluorescence to voltage, and the worm neural network, which results in a biologically more plausible computational model and enables dimensioned output and the ability to predict the polarity of chemical synapses and model parameters. Specifically, the temporal-filter module is integrated into the inference network, allowing for independent voltage predictions for each neuron based on its historical fluorescence trace. Moreover, prior knowledge is introduced into the worn neural network to enforce constraint on its biological properties. Within the framework, the chemosensory stimulus and fluorescence trace serve as inputs to the inference network, which outputs an approximation of the voltage posterior. The posterior is then sampled to obtain voltages for all neurons at each time point, and these are subsequently fed into the reconstruction model to reconstruct fluorescence. This reconstructed trace is then compared to the observed trace to calculate the reconstruction loss. Simultaneously, the worm neural network generates the voltage for the subsequent time step as a voltage prior, using inputs from the sensory input and inferred voltage. The gap between the voltage prior and posterior is quantified using the KL divergence, which, when combined with the reconstruction loss, forms the negative Evidence Lower Bound (ELBO) used for optimizing the overall model.

and the fluorescence trace **f**. It is assumed that the voltage distributions of all neurons at any given time follow a Gaussian distribution, denoting the voltage posterior for neuron i as $v_i(t) \sim \mathcal{N}(\mu_i(t), \sigma_i(t))$ for neuron $i = 1, \cdots, N$ at time point $t = 1, \cdots, D$. Consequently, the approximate posterior distribution $Q(\mathbf{v}|\mathbf{f}, \mathbf{o})$ can be derived as follows

$$Q(\mathbf{v}|\mathbf{f}, \mathbf{o}) = \prod_{t=1,\cdots,D} Q(\mathbf{v}(t)|\mathbf{f}(1), \cdots, \mathbf{f}(D), \mathbf{o}(1), \cdots, \mathbf{o}(D)) \tag{2.1}$$

where D represents the time duration of the data inputted into the inference network.

2.1.1.1. Inference network in CCLVM. Mi et al. [4] have proposed an inference network comprised of multiple spatial 1D convolutional layers (Fig. 2(a)). This network takes as input a fluorescence trace matrix **f** of size $N \times T$, where N represents the number of neurons in the given connectome and T denotes the number of time points in the trace. In cases where values are missing, they are assigned a value of zero. Along with the fluorescence trace matrix, a binary mask of size $N \times T$ is introduced, referred to as the 'missing mask', indicating the presence or absence of fluorescence data. Both the fluorescence trace and the missing mask undergo processing through an array of 1D convolutional layers, incorporating Rectified Linear Unit (ReLU) nonlinearities and upsampling layers, which result in an intermediate representation. The resulting outputs then undergo upsampling and are processed through another 1D convolutional layer. This sequence is followed by an additional ReLU nonlinearity and another layer of upsampling.

Besides the fluorescence data, a chemosensory stimulus **o** is represented as an $N \times T'$ matrix, where T' denotes the number of time points for sensory input, and is processed similarly to the fluorescence trace and missing mask. This procedure yields an intermediate representation, which is then further processed through two separate 1D convolutional

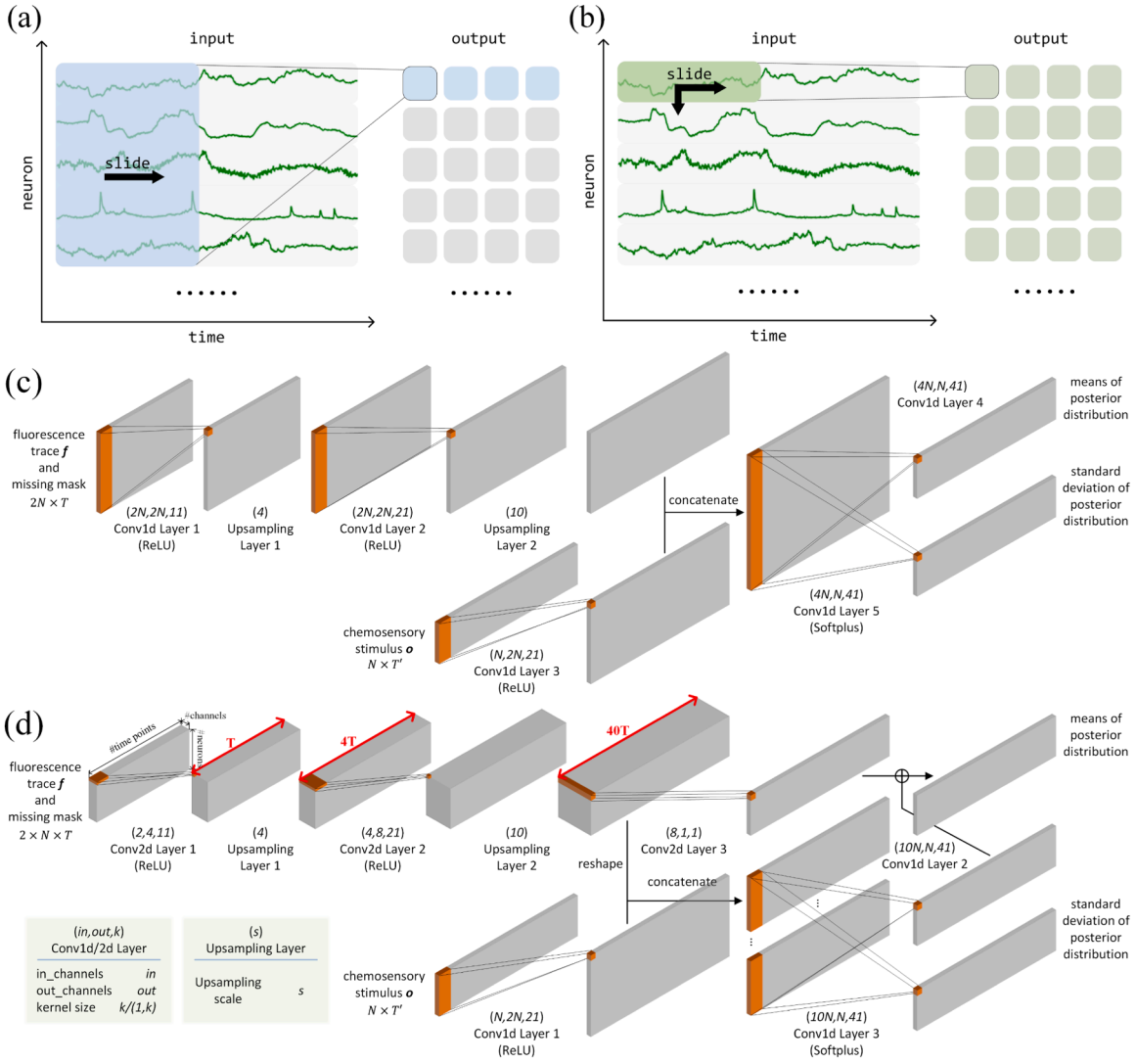


Fig. 2. Schematic diagram of the fundamental module in the inference network. (a) The 1D convolutional module utilized in CCLVM performs convolution operations on the collected fluorescence data from all neurons. Its objective is to predict each neuron's electrical activity by using distinct convolution kernels for each output neuron. (b) The temporal-filter in temporal convolutional module. This module generates predictions exclusively based on the historical signals of each individual neuron, employing a shared convolution kernel across all output neurons. (c, d) The detailed architecture of inference network. Inference network input the fluorescence trace with its missing mask and chemosensory stimulus to infer means and standard deviation of posterior Gaussian distribution. The numbers in the brackets above the labels of all network layers in the figure are hyperparameters, which are explained in the lower left corner of (d), and the activation function used by each layer is in the lower brackets. (c) The inference network of CCLVM consists of multiple spatial convolutional layers and upsampling layer. (d) Multiple temporal convolutional layers with kernel size of $1 \times k$ are applied to replace spatial convolutional layers which extract features from calcium fluorescence in TF-CCLVM and TF-PALVM. Thus, the output features of each temporal convolutional layer are obtained from the fluorescence data of their respective corresponding neurons by performing the same transformation, which aims to enhance common features across neurons.

layers. One of these layers generates the means of the posterior distribution, while the other layer produces the standard deviation, incorporating an additional softplus nonlinearity.

In the multi-channel 1D convolutional layers of this model, all channels (neurons) will be simultaneously processed to predict each neuron. However, it is noteworthy that, it is not necessary for each neuron to receive information from all other neurons to make predictions about their individual signals, given the independent nature of C. elegans [8–10]. Furthermore, during the gradient descent training process, a fully connected layer may potentially lead to the convergence at local optima.

2.1.1.2. Temporal-filter inference network. To ameliorate the aforementioned issues, we put forth a novel approach called the Temporal-Filter Inference Network (TFIN), which utilizes temporal convolutional layers equipped with $1 \times k$ (neuron dimension $\times$ time dimension) kernels (Fig. 2(b)). This approach ensures that each neuron predominantly focuses on its intrinsic features, prompting independent coding and enhancing predictive accuracy.

In TFIN, the fluorescence trace $\mathbf{f}$ and the missing mask are concatenated to form a $2 \times N \times T$ tensor. This tensor undergoes a sequence of operations: it is passed through a temporal convolutional layer, followed by a ReLU nonlinearity, an upsampling layer, another temporal convolutional layer, another ReLU nonlinearity, another upsampling layer, and finally, another temporal convolutional layer.

Concurrently, the chemosensory stimulus $\mathbf{o}$ is input into a 1D convolutional layer, followed by concatenation with the reshaped output of the second upsampling layer. The resulting combined matrix is then directed into two separate 1D convolutional layers. The output from one of these layers is added to the output of the third temporal convolutional layer, creating the means of the posterior distribution. Meanwhile, the other layer produces the standard deviation, incorporating an additional softplus nonlinearity.

It is important to note that all temporal convolutional layers in our TFIN have a kernel size of $1 \times k$ $(k = 11,21,1)$. This design ensures that the output of temporal convolutional layers for each neuron is totally derived from its unique fluorescence trace, promoting self-focused attention for each neuron. The final 1D convolutional layer is reserved with the objective of blending the temporal-filter output to predict signals for missing neurons. Through rigorous empirical evaluations, we demonstrate that TFIN significantly improves worm holdout evaluation performance.

2.1.2. Reconstruction model

The reconstruction model is empirically robust and enables the generation of reconstructed fluorescence $\widehat{\mathbf{f}}$, as a function of voltage $\mathbf{v}$, which can be described as $P(\mathbf{f}|\mathbf{v},\mathbf{o})$. Initially, voltage $\mathbf{v}$ pertaining to all neurons and temporal points, is sampled from the posterior distribution derived from the inference network $Q(\mathbf{v}|\mathbf{f},\mathbf{o})$. Subsequently, a model of calcium ion channel dynamics is applied to generate calcium concentration $[Ca^{2+}]$, as a function of the sampled voltage $\mathbf{v}$. This mathematical model is defined as follows

$$\tau_{[Ca^{2+}]}\frac{d[Ca^{2+}]_i(t)}{dt} + [Ca^{2+}]_i(t) = \text{Softplus}(v_i(t)) \tag{2.2}$$

where $\tau_{[Ca^{2+}]}$ denotes time constant of calcium leakage for each individual neuron, $[Ca^{2+}]_i(t)$ and $\frac{d[Ca^{2+}]_i(t)}{dt}$ represent the calcium concentration and its derivative for neuron i at time point t respectively, and $\text{Softplus}(x) = \log(1 + \exp(x))$.

Subsequently, an affine transformation model is employed to reconstruct the fluorescence distribution $\widehat{\mathbf{f}}$, based on the calcium concentration $[Ca^{2+}]$. This model is represented as:

$$\widehat{f}_i(t) = \alpha_i^f [Ca^{2+}]_i(t) + \beta_i^f + \sigma_i^f \epsilon_i^f(t) \tag{2.3}$$

where α_i^f, β_i^f are trainable parameters that enable the transformation from calcium concentration to fluorescence, σ_i^f represents noise amplitude, $\epsilon_i^f(t) \sim \mathcal{N}(0,1)$, and $\widehat{f}_i(t)$ represents reconstructed fluorescence distribution for neuron i at time point t, respectively.

2.1.3. Worm neural network

The worm neural network is delineated by the generative model $P(\mathbf{v}|\mathbf{o})$. According to Bargmann's observation [15] that the majority of neurons in C. elegans do not exhibit spiking behavior, a unified single-compartment model is employed to simulate the voltage dynamics for each neuron. The generative model $P(\mathbf{v}|\mathbf{o})$ describing the voltage $v_i(t)$ is given by

$$\tau_i\frac{dv_i(t)}{dt} + v_i(t) = v_i^{rest} + s_i^c(t) + s_i^e(t) + o_i(t) + \sigma_i^v \epsilon_i^v(t) \tag{2.4}$$

where τ_i is voltage time constant, v_i^{rest} is resting membrane potential, $s_i^c(t)$, $s_i^e(t)$ and $o_i(t)$ stand for chemical synaptic input, electrical synaptic input and stimulation input, respectively. σ_i^v denotes the noise amplitude, and the standard normal noise term is defined as $\epsilon_i^v(t) \sim \mathcal{N}(0,1)$. Only sensory neurons receive a non-zero stimulation input $o_i(t)$, which is generated by a nonlinear mapping from three binary sequences.

In order to model the chemical synaptic input, a conductance-based synapse model is employed. This model assumes that the quantity of neurotransmitter released is directly proportional to the pre-synaptic voltage v_j and is subject to a softplus activation. While the current-based model is also discussed in CCLVM, this study does not empirically test it, given previous research [4] demonstrating its inferiority compared to the conductance-based model. Therefore, the synaptic input $s_i^c(t)$ to a post-synaptic neuron i can be described as

$$s_i^c(t) = \sum_{j=1}^{N}(E_{ji} - v_i(t))W_{ji}^c \text{Softplus}(v_j(t)) \tag{2.5}$$

where W_{ji}^c represents the weight of the chemical synapse from neuron j to i, and E_{ji} denotes the reversal potential. Electrical synaptic input is straightforwardly defined as

$$s_i^e(t) = \sum_{j=1}^{N} W_{ji}^e(v_j(t) - v_i(t)) \tag{2.6}$$

where W_{ji}^e represents the weight of gap junction between neuron j and i.

2.1.4. Connectome-constrained network

The connectome-constrained network refers to the worm neural network in the context of CCLVM. This model operates by incorporating the anatomical connectivity derived from an actual nematode as a constraint on the neuron connectivity within the artificial worm neural network. This approach facilitates the creation of a biologically plausible model, merging a computational biological model with genuine connectivity data. It is achieved by employing synaptic connectivity matrices, denoted as $\mathbf{C}^c$ for chemical and $\mathbf{C}^e$ for electrical, as published in prior study [5]. Within this network, the weights assigned between neurons are determined in proportion to the number of synapses present in the anatomical connection.

$$\mathbf{W}^c = \alpha^c \mathbf{C}^c, \quad \mathbf{W}^e = \alpha^e \mathbf{C}^e \tag{2.7}$$

where α^c and α^e denote scale factors which can be trained respectively.

2.1.5. Physiology-aligned network

In order to represent the neural network of the nematode more accurately, we introduce the concept of physiology-aligned network. In this model, we replace the connectivity data from the connectome-constrained network [5] with recently published, more comprehensive

network connectivity data [14]. Consistent with the connectome-constrained network, the total connectivity of chemical and electrical synapses for the EM serial sections, which is a function of both the quantity and sizes of synapses, is denoted by matrices $\mathbf{A}^c$ and $\mathbf{A}^e$. The weights are established to maintain proportionality to these synapse matrices as

$$\mathbf{W}^c = \alpha^c \mathbf{A}^c, \quad \mathbf{W}^e = \alpha^e \mathbf{A}^e \tag{2.8}$$

Moreover, to better emulate an actual nematode, we assign units to variables in our model and incorporate prior knowledge about synaptic polarity and reasonable parameter ranges. We assume that all chemical synapses between neurons can be modeled with one inhibitory and one excitatory synapse, where the sum of the weights of the two synapses is equal to the original weight of all synapses, and let $\mathbf{P}^e \in [0,1]^{N\times N}$ be the weight proportion matrix of excitatory synapse to all synapses. The synaptic reversal potentials of the excitatory and inhibitory synapses are set as two trainable parameters, E_{exc} and E_{inh}, respectively. Hence, we can rewrite the conductance-based synapse model as follows

$$s_i^c(t) = \sum_{j=1}^{N} ((E_{exc} - v_i(t))P_{ji}^e + (E_{inh} - v_i(t))(1 - P_{ji}^e))W_{ji}^c \mathrm{Softplus}(v_j(t)) \tag{2.9}$$

To confine the proportion matrix $\mathbf{P}$, we leverage the chemical synapse polarity data from Fenyves et al. [11]. In their study, they estimated the partial polarities of ionotropic chemical synapses within the neuronal connectome network [14] of C. elegans based on the neurotransmitter of pre-synaptic neuron and the receptor expression of the post-synaptic neuron. Their analysis has predicted 403 synapses to be inhibitory, 1141 synapses to be excitatory, while the other 2163 synapses remain undefined. For a chemical synaptic connection from neuron j to neuron i, P_{ji} is equal to 1 if it is predicted to be excitatory, and to 0 when predicted to be inhibitory. For those synapses with undefined prediction, the excitatory proportion remains trainable.

Specifically, to ensure manageable numerical values within the neural network, we standardize the voltage unit to 10 millivolts and establish the time unit as seconds. According to Kunert et al. [13], we initialize $E_{exc} = 0\text{mV}$ and $E_{inh} = -45\text{mV}$. For each neuron i, the resting potential v_i^{rest} and time constant τ_i are initialized through a Gaussian distribution to conform $v_i^{rest} \sim \mathcal{N}(-3.5, 0.1)$ and $\tau_i \sim \mathcal{N}(0.1, 0.1)$.

2.2. Loss function

The framework's loss function consists of the reconstruction loss between observed and reconstructed fluorescence traces, as well as the Kullback-Leibler (KL) divergence between the voltage produced by the inference network and the worm neural network. This loss function is commonly referred to as the negative evidence lower bound (ELBO) loss [4]:

$$\mathscr{L} = \mathscr{L}_{recon} + D_{\mathrm{KL}} \tag{2.10}$$

The reconstruction loss $\mathscr{L}_{recon}$ is computed as the negative log likelihood between the measured and reconstructed fluorescence:

$$\begin{aligned}\mathscr{L}_{recon} &= -\mathbb{E}_{\mathbf{v}\sim Q(\mathbf{v}|\mathbf{f},\mathbf{o})}[\log(P(\mathbf{f}|\mathbf{v},\mathbf{o}))] \\ &= -\sum_{t=0}^{D\Delta t}\sum_{i=1}^{N}(\log(\sqrt{2\pi}\sigma_i^f) - \frac{(f_i(t) - \widehat{f}_i(t))^2}{2(\sigma_i^f)^2})\end{aligned} \tag{2.11}$$

where σ_i^f is standard deviation of fluorescence trace of neuron i.

D_{KL} represents the KL divergence, which quantifies the distance between the voltage prior and posterior distribution:

$$\begin{aligned}D_{\mathrm{KL}} &= D_{\mathrm{KL}}(Q(\mathbf{v}|\mathbf{f},\mathbf{o})||P(\mathbf{v}|\mathbf{o})) \\ &= \sum_{t=1}^{D\Delta t}\sum_{i=1}^{N}(\log(\frac{\sigma_i^v}{\sigma_i(t)}) + \frac{(\sigma_i(t))^2 + (\mu_i(t) - \widehat{v}_i(t))^2}{2(\sigma_i^v)^2} - \frac{1}{2})\end{aligned} \tag{2.12}$$

where σ_i^v is standard deviation of generated voltage of neuron i.

2.3. Datasets

We utilize the NeuroPAL calcium imaging dataset [3], which records the activity of neurons throughout the entire brain in immobilized worms exposed to three different chemosensory stimuli (2-butanone, 2, 3-pentanedione, and NaCl). The dataset contains the recorded activity of approximately 170 labeled neurons for each of the 21 worms, with a sampling rate of 4 Hz and a duration exceeding 3 min. Furthermore, to bolster the precision of the biophysically accurate network, we integrate the nematode connectome [14] into the physiology-aligned network. This updated connectome contains 3707 chemical and 1101 electrical connections, providing a more comprehensive representation compared to the connectome [5] used in CCLVM [4]. Additionally, we integrate prior knowledge into the physiology-aligned network by incorporating synaptic polarity data [11] and adopting a reasonable range of parameters based on previous study [13].

2.4. Metrics

To assess how well the models perform, we compute correlation coefficient r_i between recorded fluorescence $f_i(t)$ and reconstructed fluorescence $\widehat{f}_i(t)$. Correlation coefficient r_i is defined as

$$r_i = \frac{\sum_{t=1}^{D\Delta t}(f_i(t) - \bar{f}_i(t))\,(\widehat{f}_i(t) - \bar{\widehat{f}}_i(t))}{\sqrt{\sum_{t=1}^{D\Delta t}(f_i(t) - \bar{f}_i(t))^2 \sum_{t=1}^{D\Delta t}(\widehat{f}_i(t) - \bar{\widehat{f}}_i(t))^2}} \tag{2.13}$$

Moreover, mean square error (MSE) is also used to measure the absolute error between the prediction and the target. mse_i is computed as

$$mse_i = \frac{1}{D}\sum_{t=1}^{D\Delta t}(f_i(t) - \widehat{f}_i(t))^2 \tag{2.14}$$

It can be assumed that the predicted voltage is precise when the reconstructed calcium fluorescence closely resembles the recorded data, since the reconstruction model is designed to be a realistic generative model.

2.5. Statistical analysis

Each model was trained four times with different random initializations. For the correlation metric, we report the mean and 95 % confidence interval (CI) across runs. Statistical significance between TF-PALVM and CCLVM was assessed using paired two-tailed t-tests. Effect sizes were quantified using Cohen's d.

3. Results

Our study assesses the efficiency of our proposed enhancements through evaluations conducted on CCLVM and TF-PALVM. These evaluations employ both worm holdout and neuron holdout approaches, and the TF-PALVM outperforms CCLVM (conductance-based synaptic model with connectome count constraint) [4] significantly in both worm holdout and neuron holdout evaluations.

3.1. Holdout evaluations

3.1.1. Worm holdout evaluation

To evaluate the accuracy and generalization of our fluorescence-to-voltage conversion, we conduct worm holdout evaluations. Models are trained with data from only 15 out of the 21 available worms and are subsequently applied to reconstruct the fluorescence in the remaining 6 worms' data. In this scenario, the first 6 worms are designated as the testing set, while the latter 15 constitute the training set. Each model type undergoes training 4 times, with each instance employing a unique random initialization.

Fig. 3(a) presents violin plots comparing the distribution of correlation coefficients between the predicted and observed fluorescence across 2 different models, further categorized by sensory, inter, motor, and pharyngeal neurons. Additionally, we display the mean and standard error of each metric for each model, calculated based on the results of 4 random initializations (Table 2). The reported metrics encompass the mean correlation coefficient, mean square error for all recorded neurons in the first 6 worms, as well as the values of loss functions (ELBO, reconstruction loss, and KL divergence) for each model.

The inclusion of the temporal-filter inference network has resulted in substantial improvements in both correlation and MSE values when compared to the results obtained from the CCLVM. As displayed in Table 2, for TF-PALVM, the correlation increases significantly from 0.460 to 0.805, while the MSE decreases from 0.292 to 0.163. The predicted traces depicted in Fig. 3(b) vividly demonstrate the enhanced accuracy of the model employing the temporal-filter inference network in reconstructing calcium fluorescence data. These findings furnish compelling evidence for the effectiveness of the temporal convolutional layer in accurately converting calcium fluorescence data into voltage signals.

Overall, TF-PALVM achieves a mean correlation of 0.805 (95 % CI: [0.778, 0.833]) compared to 0.460 (95 % CI: [0.413, 0.506]) for CCLVM, corresponding to a ~75 % relative improvement. The difference is statistically significant (paired t-test, $p = 1.4 \times 10^{-5}$, Cohen's d=27.1).

Table 2
Results for holdout evaluations. The table displays the mean and standard error (SE) of each metric for each model, trained with four different random initializations, in both the worm holdout evaluation and neuron holdout evaluation. Overall, TF-PALVM outperforms CCLVM in all metrics. Results that perform better than CCLVM are bolded. ↑(↓) indicates that larger (smaller) value is preferable.

holdout	metric		model		
			CCLVM [4]	CCLVM*	TF-PALVM
worm	*corr*↑	mean	0.474	0.460	**0.805**
		SE	±0.006	±0.013	**±0.007**
	MSE↓	mean	-	0.292	**0.163**
		SE	-	±0.008	**±0.005**
	ELBO↑	mean	-0.854	-0.868	**-0.671**
		SE	±0.036	±0.008	**±0.009**
	recon↓	mean	0.744	0.820	**0.635**
		SE	±0.035	±0.009	**±0.010**
	KLD↓	mean	0.110	0.048	**0.036**
		SE	±0.006	±0.001	**±0.002**
neuron	*corr*↑	mean	0.318	0.282	**0.425**
		SE	±0.016	±0.011	**±0.002**
	MSE↓	mean	-	3.170	**1.528**
		SE	-	±0.017	**±0.048**

* We reproduce the CCLVM results based on previous study [4].

3.1.2. Neuron holdout evaluation

To assess the model's capability to accurately predict the activity of neurons that have not been previously observed, we perform a neuron holdout evaluation. In this evaluation, the calcium fluorescence data for a single neuron or a pair of bilaterally symmetric neurons is omitted from both the training and testing sets. Owing to the ambiguity in the original paper by Mi et al. [4] regarding specific worm used for training and testing, we executed the neuron holdout evaluation on worm 1, which shares an equal number of 170 recorded neurons with the CCLVM

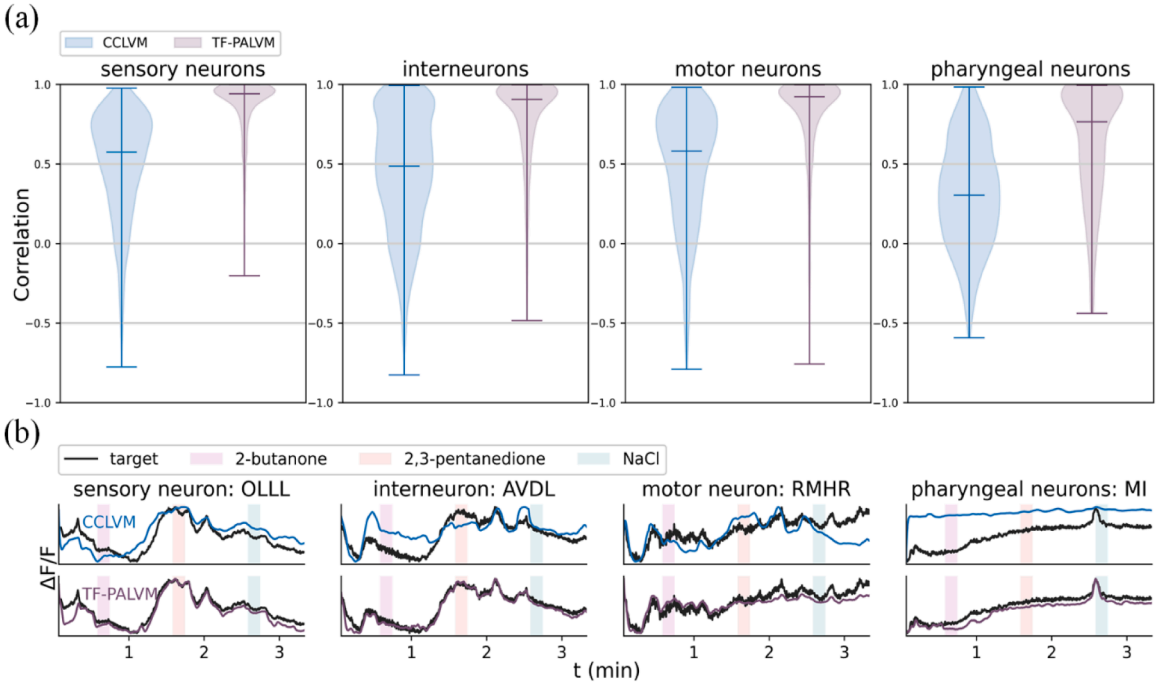


Fig. 3. Reconstruction results for different neuron categories in worm holdout evaluation. (a) displays a comparison of correlation coefficient distributions of reconstructed and measured fluorescence for two different models in a violin plot. The results of sensory, inter, motor and pharyngeal neurons are shown separately. (b) shows the measured and predicted fluorescence by the two different models for four chosen neurons: sensory neuron OLLL, interneuron AVDL, motor neuron RMHR and pharyngeal neuron MI.

test.

Fig. 4 showcases violin plots comparing the distribution of correlation coefficients between predicted and measured fluorescence, along with the measured and predicted fluorescence traces for four specifically chosen neurons, for each model under the neuron holdout evaluation. Considering that the initial calcium fluorescence is undetermined for the neurons held out, the results of the first 30 points, equivalent to approximately 8 s, are omitted from the trace plot. Table 2 enumerates the correlation and mean square error (MSE) for each model, computed based on the results of four independent random initializations.

The evolution from CCLVM to TF-PALVM is accompanied by a notable enhancement in correlation, increasing from 0.282 to 0.425. Concurrently, there is a significant reduction in MSE, decreasing from 3.170 to 1.528. These results emphasize that the integration of prior knowledge and units can make the generative model more biologically plausible, enabling it to predict neuronal activity that closely resembles reality.

TF-PALVM outperforms CCLVM with a mean correlation of 0.425 (95 % CI: [0.417, 0.432]) versus 0.282 (95 % CI: [0.249, 0.316]), a ~51 % improvement (paired t-test, $p = 0.0012$, Cohen's d=6.12).

3.2. Unconstrained retraining

To further explore the realism of synaptic weights in generative models, we conducted retraining of both CCLVM and TF-PALVM, removing the connectome constraint. This approach aligns with the variant of CCLVM referred to as "connectome count2" in the study by Mi et al. [4]. This process involves initializing the model with the trained parameters, allowing both the chemical and electrical synapse weights $\mathbf{W}^c, \mathbf{W}^e$ to be learnable, and retraining the model. Each model is trained four times with different random initialization. To ensure the uniqueness of the predicted connectome, only data from worm 1 are used for training, with no neurons being held out during this process.

The distribution of reversal potentials, synaptic polarity, and neuron time constants for both models are depicted in Fig. 5(a-c), and predictions from TF-PALVM regarding the proportion of chemical synapse polarity are presented in Fig. 5(d). Additionally, we provide polarity ratio predictions for all chemical synapses, including those not predicted by Fenyves et al. [11], in the supplementary file: chem_sign.xls.

Significantly, the prediction from TF-PALVM for the excitatory ratio in chemical synapses (74.63 %) closely aligns with the ratio obtained through neurotransmitter and receptor matching (76.88 %), whereas the CCLVM prediction (48.88 %) exhibits a considerable deviation (Fig. 5(b)). Additionally, it is important to highlight that while Fenyves et al. [11] are able to predict approximately 48 % of connections, specifically those that are purely excitatory or inhibitory, our model successfully makes predictions for all connections and can predict the ratio of excitatory and inhibitory synapses for each neuron interconnection. This result unequivocally demonstrates that parameter in TF-PALVM closely aligns with the evidence from real worm physiology.

3.3. Ablation studies

3.3.1. Input ablation

To investigate the importance of different neurons in predicting neural activity, we systematically perform input ablation utilizing well-trained TF-PALVMs in neuron holdout evaluation. This involves ablating the fluorescence signals of pairs of neurons from the input and subsequently calculating the average correlation between the reconstructed and measured fluorescence data when predicting various held-out neurons.

In Fig. 6(a), we present the sorted neuron pairs based on their mean correlation following input ablation, and plot the resulting correlation curve. Our findings demonstrate that the ablation of hub neurons (AVA, AVB) [18] has a substantial influence on correlation. Additionally, among the 96 pairs of neurons with available data, we identify the top 5 pairs, namely RMDV, AVA, RME, AVB and ASE, which exerts the greatest impact on correlation. We have observed that these neurons likely correspond to distinct functional modules within the neural network of C. elegans. The RMD neurons, recognized as central pattern generator (CPG), play a crucial role in driving locomotion in the nematode. The interneurons AVB (AVA) are key components involved in

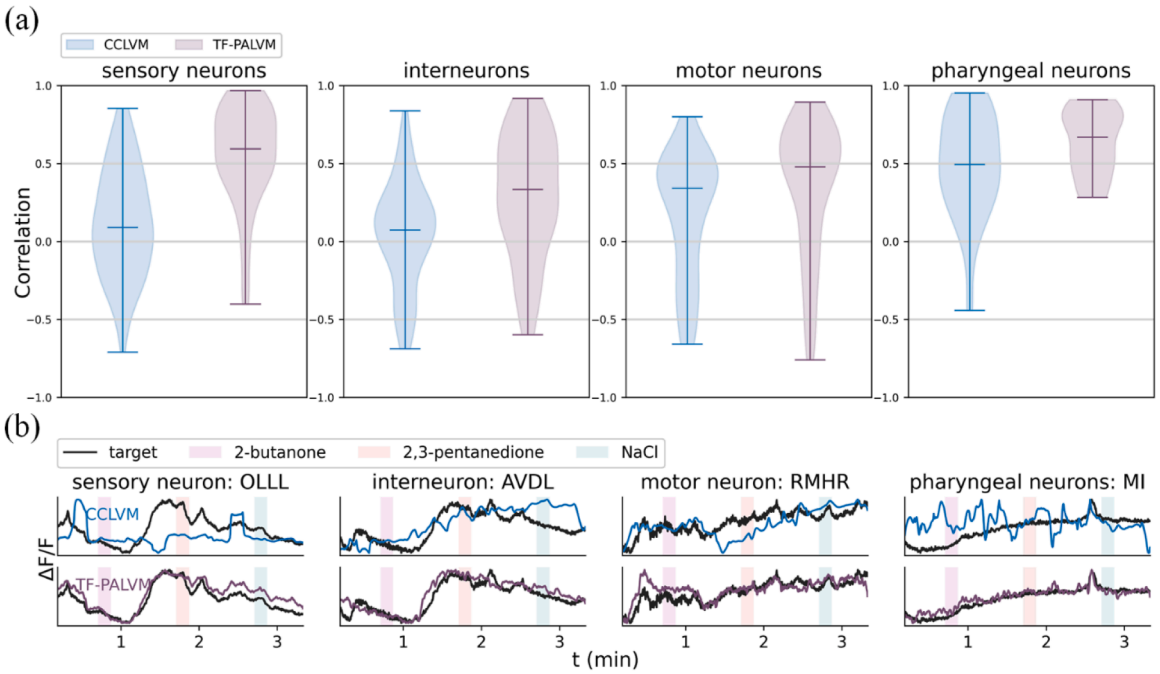


Fig. 4. Prediction results for different neuron categories on neuron holdout evaluation. (a) displays a comparison of correlation coefficient distributions of predicted and measured fluorescence for two different models in a violin plot. The results of sensory, inter, motor and pharyngeal neurons are shown separately. (b) shows the measured and predicted fluorescence by the 2 different models for four chosen neurons: sensory neuron OLLL, interneuron AVDL, motor neuron RMHR and pharyngeal neuron MI.

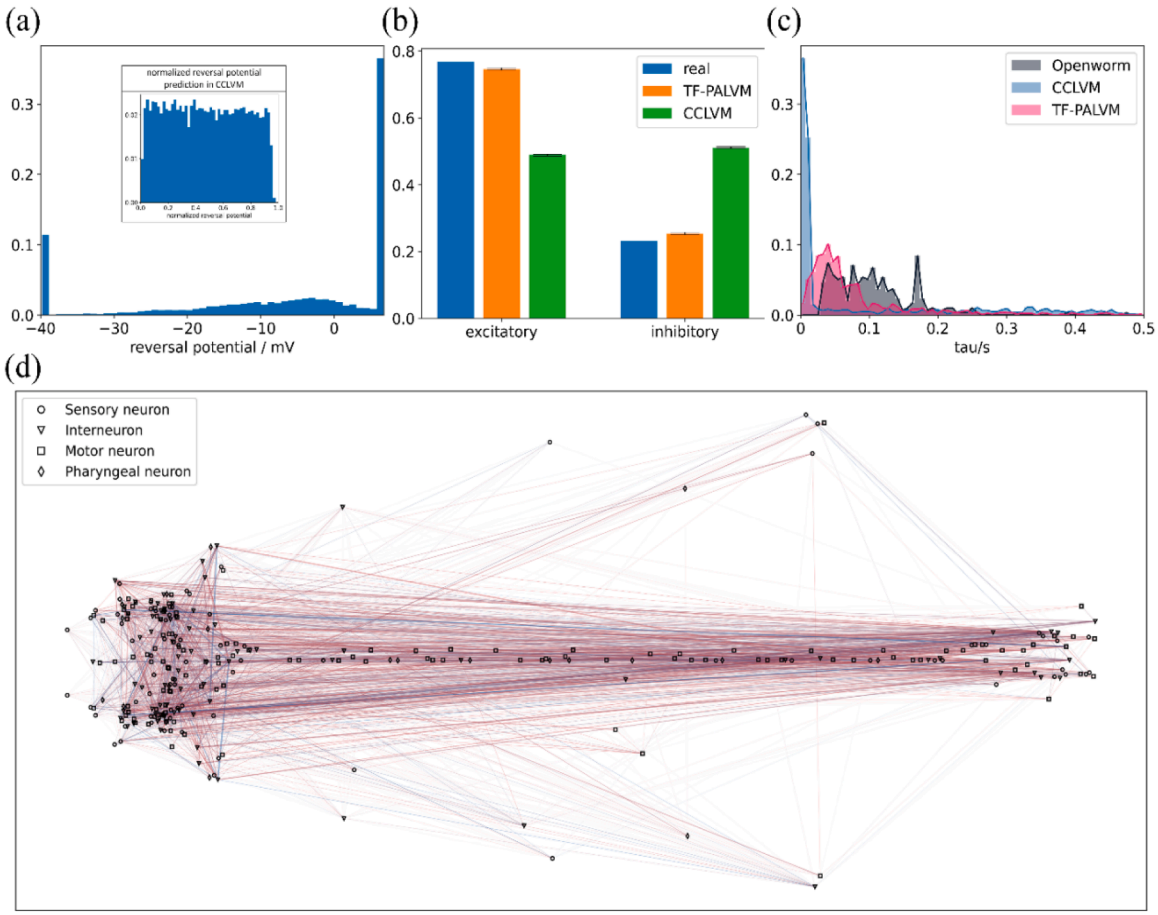


Fig. 5. Prediction for reversal potential, synaptic polarity and time constant. (a) The distribution of reversal potentials with unit in TF-PALVM and dimensionless reversal potentials in CCLVM. Nearly half of the neuronal connections are predicted to be purely inhibitory or excitatory, while the rest are predicted to have both inhibitory and excitatory synapses. The higher the proportion of excitatory synapses, the higher the reversal voltage, and vice versa. (b) Comparison of the ratios of excitatory and inhibitory synapses predicted by experimental measurements [11], TF-PALVM, and CCLVM models. TF-PALVM and CCLVM models provide predicted excitatory-inhibitory synapse ratios for all neuron connections, while experimental results are obtained by measuring a subset (~48 %) of neuron connections, which are purely inhibitory or excitatory. (c) The distribution of neuron time constants, as compared across predictions from TF-PALVM, CCLVM, and time constant measurements at OpenWorm [16,17] based on the C2 parameter set, is presented. The distribution of time constants in CCLVM shows that approximately 60 % of the values fall below 25 ms. In contrast, the predictions from TF-PALVM closely match the observed distribution in OpenWorm, indicating a higher level of accuracy. (d) A map of synaptic polarity distribution of neural network connections in C. elegans. Circles, triangles, squares and rhombuses represent sensory, inter, motor, pharyngeal neurons respectively. The color of the line between neurons indicates the ratio of synaptic polarity, with redder lines indicating higher excitatory ratio and bluer lines indicating higher inhibitory ratio. The thickness of the line represents the connection weight of the synapse, with thicker lines indicating stronger connections.

forward (backward) movement [19]. The RME neurons influence head bending amplitude [20], and its elimination results in a distinct looping behavior confined to the head region [21]. The ASE neurons serve as primary responding neurons in the calcium fluorescence dataset NeuroPAL, which presents three types of stimuli [3] and ASER has the capability to memorize the chemotactic set point [22].

By examining the aforementioned categories of behavior-affecting neurons, we have observed that each pair of neurons among the top 5 pairs correspond to a specific class. This suggests that these neurons exert a more significant influence within their respective categories of neurons. Furthermore, despite the relatively subdued impact of ablations on RIV and RIM, it is also noteworthy that stimulation of RIV motor neurons induces omega turns and RIM neurons are considered navigation circuits that connect with AVA and AVB. Distinctive background colors are utilized to highlight different types of neurons on the horizontal axis (Fig. 6(a)).

Additionally, we have systematically ablated each pair of neurons, starting from the most influential and proceeding towards the least influential, as well as in the opposite direction. Fig. 6(b) illustrates the changes in correlation, with a gray line denoting the 90 % correlation threshold when no neuron is ablated. Our results suggest that eliminating the two most influential neuron pairs results in a notable reduction in correlation, bringing it down below 90 % of the original accuracy. However, when the least influential neurons are removed, it necessitates the ablation of 86 neuron pairs to reach a correlation of 90 %. Put differently, achieving a prediction efficacy of 90 % can be attained by preserving only the 11 most influential neuron pairs, implying that a limited number of key neurons play a pivotal role in predicting the activity of all neurons. This observation sheds light on the independent encoding nature of neuronal activity in C. elegans.

3.3.2. *Feature ablation*

To further elucidate the independent coding capability of the temporal-filter inference network, we conduct a feature ablation experiment using well-trained TF-PALVMs and CCLVMs in the worm holdout evaluation.

We perform feature ablation on the final 1D convolutional layer of both the TF-PALVMs and CCLVMs. Since each feature corresponds to a specific neuron in TF-PALVM, we conduct experiments to ablate the features of either the neuron itself or other neurons. Conversely, in

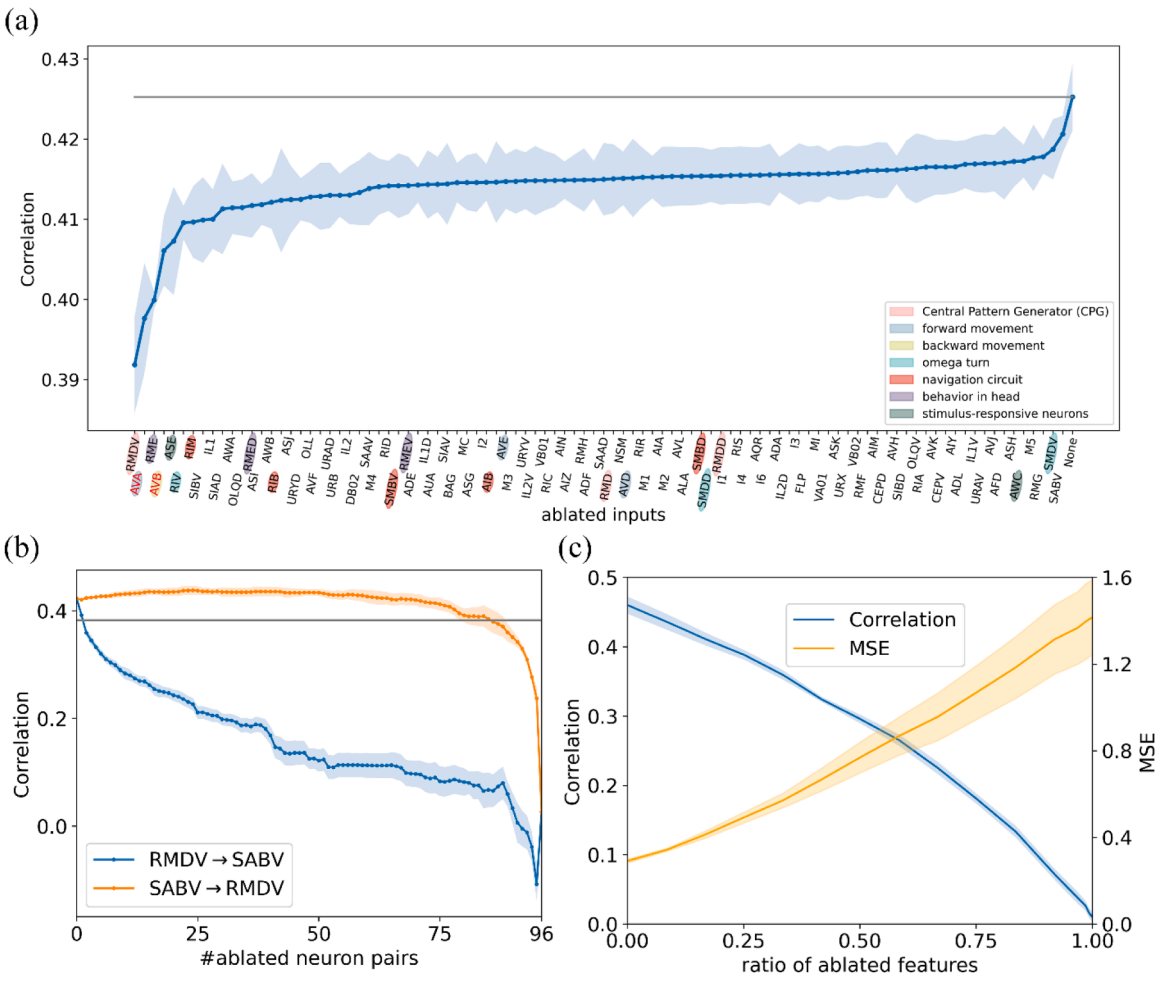


Fig. 6. Ablation study for different neurons and features. (a) The correlation curve predicted by neuron holdout evaluation after ablating the fluorescence of each pair of neurons, with the order of the neuron pairs sorted according to mean correlation. Hub neurons (AVA and AVB) [18] are marked in red. The gray line corresponds to the correlation baseline when no input is ablated, indicated as "None". Distinctive background colors are utilized to highlight different types of neurons on the horizontal axis. Ablating the leftmost neurons has the most significant impact on predicting neuron activity, suggesting their more crucial role within their respective functional modules of neurons. (b) The correlation curve calculated by neuron holdout evaluation by deleting each pair of neurons in turn from RMDV to SABV (blue) or SABV to RMDV (orange) in (a). The gray line represents correlation$= 0.9 \times 0.425$. (c) The curves depicting the correlation and MSE of CCLVM results demonstrate a smooth change as features ablated, indicating a dense relationship between features and predictions.

CCLVM, where the features are unordered, we employ different ratios for ablation.

Specifically, during the reconstruction of a specific neuron in TF-PALVM, we selectively ablate the features associated with that neuron, accounting for approximately 0.33 % of the total features. As a result, the correlation decreases significantly from 0.805 to 0.273. However, when we carry out ablation on all other features unrelated to that neuron, constituting approximately 99.67 % of the total features, the correlation increases to 0.876. This observation highlights the independent coding property inherent in the temporal-filter inference network.

In the case of CCLVM, each feature exhibits a dense correlation with all neurons. To investigate the effects of feature ablation, we randomly ablate various proportions of features and calculate the corresponding correlation and MSE (Fig. 6(c)). Our findings indicate that as features are incrementally ablated, both correlation and MSE change smoothly. This provides further evidence for the population coding characteristic of the CCLVM, contrary to previous experimental-based studies [9,10].

3.3.3. *Locomotion ablation*

Moreover, we conduct ablation tests on each of the 96 pairs of neurons to assess their impact on locomotion, employing the digital twin C. elegans model [23]. Subsequently, we provide a comprehensive comparison between the changes in prediction correlation and those in crawling distance and chemotaxis index, when ablating each pair of neurons (Fig. 7). Notably, greater reductions in crawling distance and chemotaxis index signify a more substantial impact after resulting from the ablation of the neuron pair. We observe that the 5 neuron pairs with the greatest impact on correlation also exerted significant effects on crawling distance and chemotaxis index. We suggest that their ablation may disrupt the overall activities of the neural population, thereby influencing locomotion. However, for the neurons located in the top left quadrant delimited by the two gray lines, their ablation had a minor effect on correlation but a substantial impact on locomotion. This indicates that the ablation of these neurons may directly affect locomotor behavior. Specifically, in Fig. 7(a), the top 5 neuron pairs in the top left quadrant include RMED, URYD, RIC, RMD, and RMG. In Fig. 7(b), the top 5 neuron pairs in the top left quadrant include ADF, AIA, RMDD, VA01, and ASH. The ablation of RMED has been found to result in an increase in head bending amplitude [20]. The circuit activity of RMG has been linked to locomotion arousal and quiescence, promoting corresponding responses to oxygen, pheromones, and food [24,25]. ADF acts as post-synaptic to ASH, which mediates chemotaxis to various chemicals and is responsible for toxin avoidance [26,27]. AIA neurons

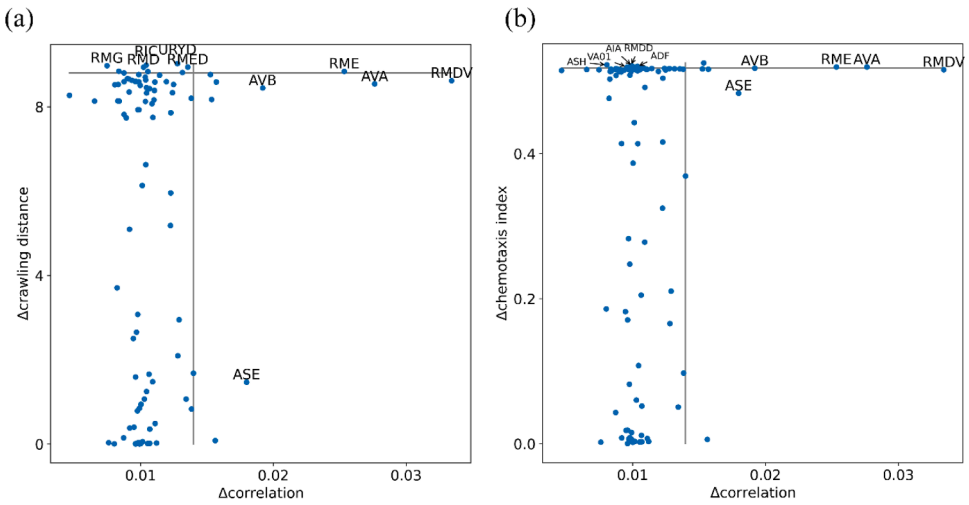


Fig. 7. Comparison between the absolute changes in correlation and those in (a) crawling distance and (b) chemotaxis index resulting from ablating different pairs of neurons. The gray line demarcates the top 10 pairs of neurons with the largest changes. The 5 pairs of neurons with the largest changes of each metrics are marked.

have been identified as contributors to chemotactic behavior [28,29]. RMD and RMDD play a crucial role in controlling the muscles in the head and neck, facilitating spontaneous foraging movements that are of significant importance for chemotaxis behavior [21,30].Additional information about the digital twin C. elegans can be found in the Supplementary. For detailed test results of all neurons, please refer to the supplementary file: locomotion_results.xlsx.

4. Discussions

In this study, we introduce innovative enhancements to the CCLVM framework [4], resulting in the TF-PALVM, which closely aligns with the real worm not only in activity but also in physiology. Unlike the dense connectivity in the CCLVM, the TF-PALVM innovatively integrates temporal convolutional layer into the inference network, aligning with the independent coding nature in C. elegans [8–10]. This addition effectively captures the historical response of individual neurons, effectuating a substantial improvement in converting calcium fluorescence to voltage. Additionally, the integration of voltage units into the worm neural network characterizes another notable innovation in our approach. This incorporation empowers us to utilize prior knowledge, including parameter values [13] and synaptic polarity [11]. Consequently, our model's biological plausibility witnesses an appreciable enhancement. Subsequent to these advancements, our TF-PALVM demonstrates substantially elevated accuracy in predicting previously unobserved neuronal activity during neuron holdout evaluations. Significantly, our results also align with experimentally derived predictions [11] regarding the proportion of excitatory synapses within the trained TF-PALVM. This concordance underscores the vital role of incorporating prior knowledge, which facilitates the accurate prediction of physiological parameters that are often challenging to measure in live worms.

Furthermore, our implementation of ablation studies and a neuron sorting analysis sheds light on the outsized contribution of specific neurons (RMDV, AVA, RME, AVB and ASE), alongside the demonstrable independent coding attributes of our temporal-filter inference network. Notably, AVA and AVB have previously been identified as hub neurons by Chen et al. [18]. Considering the similar importance of RMDV, RME, and ASE neurons in accurate prediction to that of hub neurons, we propose that these neurons likely hold comparable significance. Moreover, existing evidence suggests their vital roles in motion and perception [19,20,22].

Finally, multisensory processing and motor decision-making are critical for survival in the external world [31]. Our framework, underpinned by these novel elements, holds promising potential for expansion to assimilate a derivable muscle and body model [2], thereby facilitating the examination of locomotion behavior in C. elegans. The integration of both body posture and calcium fluorescence data [2], combined with neuronal recognition algorithms [32,33], and the inclusion of additional prior knowledge of behavior, paves the way for training a worm agent. This agent, encapsulating both neuronal activity and body movement, enables a more holistic exploration of the underpinning mechanisms involved in nematode motor control, marking a significant step forward in our understanding of complex biological systems.

4.1. Limitations

Generalization Capability: A notable limitation of our study is the unexplored generalizability of the Temporal-Filter and Physiology-Aligned Latent Variable Model (TF-PALVM) to other organisms or more complex neural systems. The current model's remarkable predictive accuracy and biological plausibility are demonstrated within the relatively simple neural system of C. elegans, comprising around 302 neurons. However, this simplicity also raises questions about the model's applicability to more complex organisms with vastly larger and more intricate neural networks. Future iterations of this model should aim to explore its adaptability and effectiveness in broader biological contexts, potentially requiring modifications to accommodate the increased complexity of larger neural systems.

Model Complexity and Computational Demand: Another aspect that warrants attention is the computational complexity and resource demands of the TF-PALVM. Our manuscript has not delved into the computational requirements or the potential challenges that might be faced by researchers with limited computational resources. The model, while robust and efficient within the scope of our study, might impose significant computational burdens when scaled up or applied to more complex datasets. It is important for future research to optimize the model's computational efficiency and explore ways to reduce its resource demands without compromising its predictive accuracy. This would make the model more accessible and practical for a wider range of researchers, especially those working in resource-constrained environments.

4.2. Reproductivity

We trained each of LVMs on 1 RTX 3090. The implementation of TF-

PALVM is publicly available at: https://github.com/chongjg/TFPALVM.

CRediT authorship contribution statement

Zhongyu Chen: Visualization, Validation. **Jianfeng Feng:** Writing – review & editing, Supervision, Funding acquisition. **Chong Li:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation. **Xiangyang Xue:** Writing – review & editing, Supervision, Funding acquisition. **Yuguo Yu:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Thanks for the support from the Science and Technology Innovation 2030 - Brain Science and Brain-Inspired Intelligence Project (2021ZD0201301 and 2021ZD0200204), the National Natural Science Foundation of China (9257020, U20A20221), the Shanghai Municipal Science and Technology Major Project (2018SHZDZX01 and 2021SHZDZX0103) and ZJLab, Shanghai Municipal Science and Technology Committee of Shanghai (24JS2810400 and 21XD1400400), Shanghai Technology Development and Entrepreneurship Platform for Neuromorphic and AI SoC and in part by the Shanghai Research and Innovation Functional Program under Grant 17DZ2260900. We thank Shanghai Institute for Mathematics and Interdisciplinary Sciences (SIMIS) for their financial support (SIMIS-ID-2025-NC). The computations in this research were performed using the CFFF platform of Fudan University.

Appendix A. Supporting information

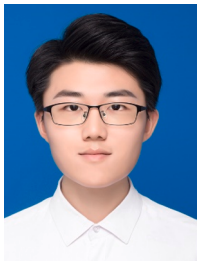
Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neucom.2026.132705.

Data availability

Data will be made available on request.

References

[1] Saul Kato, Harris S. Kaplan, Tina Schrödel, Susanne Skora, Theodore H. Lindsay, Eviatar Yemini, Shawn Lockery, Manuel Zimmer, Global brain dynamics embed the motor command sequence of caenorhabditis elegans, Cell 163 (2015) 656–669.
[2] Jeffrey P. Nguyen, Ashley N. Linder, George S. Plummer, Joshua W. Shaevitz, Andrew M. Leifer, Automatically tracking neurons in a moving and deforming brain, PLOS Comput. Biol. 13 (2017) e1005517.
[3] Eviatar Yemini, Albert Lin, Amin Nejatbakhsh, Erdem Varol, Ruoxi Sun, Gonzalo E. Mena, Aravinthan D.T. Samuel, Liam Paninski, Vivek Venkatachalam, Oliver Hobert, NeuroPAL: a multicolor atlas for whole-brain neuronal identification in C. elegans, Cell 184 (2021) 272–288, e11.
[4] Richard Xu Mi, Lu, Sridhama Prakhya, Albert Lin, Nir Shavit, Aravinthan D. T. Samuel, Srinivas C. Turaga, Connectome-constrained latent variable model of whole-brain neural activity, Int. Conf. Learn. Represent. (2022).
[5] James G. White, Eileen Southgate, J.Nichol Thomson, Sydney Brenner, The structure of the nervous system of the nematode Caenorhabditis elegans, Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 314 (1165) (1986) 1–340.
[6] Bruno B. Averbeck, Peter E. Latham, Alexandre Pouget, Neural correlations, population coding and computation, Nat. Rev. Neurosci. 7 (2006) 358–366.
[7] Sheila Nirenberg, Steve M. Carcieri, Adam L. Jacobs, Peter E. Latham, Retinal ganglion cells act largely as independent encoders, Nature 411 (2001) 698–701.
[8] Edward D. Lee, Xiaowen Chen, Bryan C. Daniels, Discovering sparse control strategies in neural activity, PLOS Comput. Biol. 18 (2022) e1010072.
[9] Kerem Uzel, Saul Kato, Manuel Zimmer, A set of hub neurons and non-local connectivity features support global brain dynamics in C. elegans, Curr. Biol. 32 (2022) 3443–3459, e8.
[10] Alon Zaslaver, Oshrat Shtangel Idan Liani, Lisa Yee Shira Ginzburg, W. Sternberg Paul, 'Hierarchical sparse coding in the sensory system of Caenorhabditis elegans', Proc. Natl. Acad. Sci. 112 (2015) 1185–1189.
[11] Bánk G. Fenyves, Gábor S. Szilágyi, Zsolt Vassy, Csaba Sőti, Peter Csermely, Synaptic polarity and sign-balance prediction using gene expression data in the Caenorhabditis elegans chemical synapse neuronal connectome network, PLOS Comput. Biol. 16 (2020) e1007974.
[12] Michael R. Harris, P.Wytock Thomas, A.Kovács István, Computational inference of synaptic polarities in neuronal networks, Adv. Sci. 9 (2022) 2104906.
[13] James Kunert, Eli Shlizerman, J.Nathan Kutz, Low-dimensional functionality of complex network dynamics: neurosensory integration in the Caenorhabditis elegans connectome, Phys. Rev. E Stat. Nonlinear Soft Matter Phys. 89 (2014) 052805.
[14] Steven J. Cook, Travis A. Jarrell, Christopher A. Brittin, Yi Wang, Adam E. Bloniarz, Maksim A. Yakovlev, Ken C.Q. Nguyen, Leo T.H. Tang, Emily A. Bayer, Janet S. Duerr, Hannes E. Bülow, Oliver Hobert, David H. Hall, Scott W. Emmons, Whole-animal connectomes of both Caenorhabditis elegans sexes, Nature 571 (2019) 63–71.
[15] C.I. Bargmann, Neurobiology of the Caenorhabditis elegans genome, Science (New York N. Y.) 282 (1998) 2028–2033.
[16] Padraig Gleeson, Radu Grosu David Lung, Ramin Hasani, Stephen D. Larson, 'c302: a multiscale framework for modelling the nervous system of Caenorhabditis elegans', Philos. Trans. R. Soc. B Biol. Sci. 373 (2018) 20170379.
[17] Gopal P. Sarma, Tom Chee Wai Lee, Vahid Portegys, Travis Ghayoomie, Bradly Jacobs, Matteo Alicea, Michael Cantarelli, Richard C. Currie, Shane Gerkin, Padraig Gingell, Richard Gleeson, Ramin M. Gordon, Giovanni Hasani, Sergey Idili, David Khayrulin, Andrey Lung, Mark Palyanov, Watts, D.Larson Stephen, 'OpenWorm: overview and recent advances in integrative biological simulation of Caenorhabditis elegans', Philos. Trans. R. Soc. B Biol. Sci. 373 (2018) 20170382.
[18] Yuhan Chen, Shengjun Wang, Claus C. Hilgetag, Changsong Zhou, Trade-off between multiple constraints enables simultaneous formation of modules and hubs in neural systems, PLOS Comput. Biol. 9 (2013) e1002937.
[19] Arisaka, Katsushi. 2022. 'Grand Unified Theory of Mind and Brain-Part I: Space-time Approach to Dynamic Connectomes of C. Elegans and Human Brains by Mepmos'.
[20] Shen, Quan Wen Yu, He Liu, Connie Zhong, Yuqi Qin, Gareth Harris, Taizo Kawano, Min Wu, Tianqi Xu, Aravinthan D.T. Samuel, Yun Zhang, An extrasynaptic GABAergic signal modulates a pattern of forward movement in Caenorhabditis elegans, eLife 5 (2016) e14197.
[21] Jesse M. Gray, Joseph J. Hill, Cornelia I. Bargmann, A circuit for navigation in Caenorhabditis elegans, Proc. Natl. Acad. Sci. 102 (2005) 3184–3191.
[22] Linjiao Luo, Quan Wen, Jing Ren, Michael Hendricks, Marc Gershow, Yuqi Qin, Joel Greenwood, Edward R. Soucy, Mason Klein, Heidi K. Smith-Parker, Ana C. Calvo, Daniel A. Colón-Ramos, Aravinthan D.T. Samuel, Yun Zhang, Dynamic encoding of perception, memory, and movement in a C. elegans chemotaxis circuit, Neuron 82 (2014) 1115–1128.
[23] Zhongyu Chen, Yuguo Yu, Xiangyang Xue, A connectome-based digital twin caenorhabditis elegans capable of intelligent sensorimotor behavior, Mathematics (2023).
[24] Heeun Jang, Sagi Levy, Steven W. Flavell, Fanny Mende, Richard Latham, Manuel Zimmer, Cornelia I. Bargmann, Dissection of neuronal gap junction circuits that regulate social behavior in Caenorhabditis elegans, Proc. Natl. Acad. Sci. 114 (2017) E1263–E1272.
[25] Seungwon Choi, Kelsey P. Taylor, Marios Chatzigeorgiou, Zhitao Hu, William R. Schafer, Joshua M. Kaplan, Sensory neurons arouse C. elegans locomotion via both glutamate and neuropeptide release, PLOS Genet. 11 (2015) e1005359.
[26] Tai-Hong Wu Guo, Min, Yan-Xue Song, Ming-Hai Ge, Chun-Ming Su, Wei-Pin Niu, Lan-Lan Li, Zi-Jing Xu, Chang-Li Ge, Maha T.H. Al-Mhanawi, Shi-Ping Wu, Zheng-Xing Wu, Reciprocal inhibition between sensory ASH and ASI neurons modulates nociception and avoidance in Caenorhabditis elegans, Nat. Commun. 6 (2015) 5655.
[27] Cornelia I. Bargmann, H.Robert Horvitz, Chemosensory neurons with overlapping functions direct chemotaxis to multiple chemicals in C. elegans, Neuron 7 (1991) 729–742.
[28] Larsch, Johannes, Steven W.Flavell, Qiang Liu, Andrew Gordus, Dirk R. Albrecht, and Cornelia I. Bargmann. 2015. 'A Circuit for Gradient Climbing in em C.& xa0; elegans em Chemotaxis', Cell Reports, 12: 1748-1760.
[29] Wang, Lifang, Hirofumi Sato, Yohsuke Satoh, Masahiro Tomioka, Hirofumi Kunitomo, and Yuichi Iino. 2017. 'A Gustatory Neural Circuit of Caenorhabditis elegans Generates Memory-Dependent Behaviors in Na+ Chemotaxis', The Journal of Neuroscience, 37: 2097-2111.
[30] Martina Nicoletti, Alessandro Loppini, Letizia Chiodo, Viola Folli, Giancarlo Ruocco, Simonetta Filippi, Biophysical modeling of C. elegans neurons: Single ion currents and whole-cell dynamics of AWCon and RMD, PLOS ONE 14 (2019) e0218738.
[31] Jiawei Zhang, Yong Gu, Aihua Chen, Yuguo Yu, Unveiling dynamic system strategies for multisensory processing: from neuronal fixed-criterion integration to population bayesian inference, Research 2022 (2022).
[32] Yuxiang Wu, Shang Wu, Xin Wang, Chengtian Lang, Quanshi Zhang, Quan Wen, Tianqi Xu, Rapid detection and recognition of whole brain activity in a freely behaving Caenorhabditis elegans, PLOS Comput. Biol. 18 (2022) e1010594.
[33] Shivesh Chaudhary, Sol Ah Lee, Yueyi Li, Dhaval S. Patel, Hang Lu, Graphical-model framework for automated annotation of cell identities in dense cellular images, eLife 10 (2021) e60321.



Chong Li received the B.S. degree in Electronic Information School from Wuhan University in 2021. He is currently pursuing a Ph.D. degree in Artificial Intelligence at College of Computer Science and Artificial Intelligence, Fudan University, under the supervision of Dr. Xiangyang Xue. His research interests include episodic memory and multi-modal modeling.

Zhongyu Chen received the B.S. and M.S. degrees in computer science from Fudan University, Shanghai, China, in 2020 and 2023, respectively. He is currently pursuing a Ph.D. degree in computer science at the College of Computer Science and Artificial Intelligence, Fudan University, under the supervision of Dr. Xiangyang Xue. His research interests include computational neuroscience and brain-inspired intelligence.

Jianfeng Feng received the BS, MS, and PhD degrees from the Department of Probability and Statistics, Peking University, China. He is the chair professor with the Shanghai National Centre for Mathematic Sciences and the dean with the Brain-Inspired AI Institute, Fudan University. He leads the DTB project. He has been developing new mathematical, statistical, and computational theories and methods to meet the challenges raised in neuroscience and mental health research.

Xiangyang Xue received the BS, MS, and PhD degrees in communication engineering from Xidian University, Xian, China, in 1989, 1992, and 1995, respectively. He is currently a professor in computer science with Fudan University, Shanghai, China. His research interests include multimedia information processing and machine learning.

Yuguo Yu is a full professor at the Research Institute of Intelligent and Complex Systems, Fudan University. He earned a Ph. D. in Physics from Nanjing University (2001), conducted postdoctoral research in Computational Neuroscience at Carnegie Mellon University (2001–2004), and served as Associate Research Scientist at Yale Medical School (2005–2012). His research focuses on neural information processing, spiking neural network modeling, neuro-energetics, and brain-inspired intelligence. He has published over 80 papers in journals including Nature, PNAS, and Physical Review Letters. He was awarded the Eastern Scholar Professorship (2013; 2017) and Shanghai Excellent Academic Leader (2021).