1

A FEED-FORWARD NEURAL NETWORK FOR INCREASING THE HOPFIELD-NETWORK STORAGE CAPACITY

Shaokai Zhao¹

College of Life Sciences, Nankai University, 300071 Tianjin, PR China

Bin Chen

College of Life Sciences, Nankai University, 300071 Tianjin, PR China

Hui Wang

College of Life Sciences, Nankai University, 300071 Tianjin, PR China

Zhiyuan Luo

Department of Computer Science, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK

Tao Zhang*

College of Life Sciences, Nankai University, 300071 Tianjin, PR China

In the hippocampal dentate gyrus (DG), pattern separation mainly depends on the concepts of 'expansion recoding', meaning random mixing of different DG input channels. However, recent advances in neurophysiology have challenged the theory of pattern separation based on these concepts. In this study, we propose a novel feed-forward neural network, inspired by the structure of the DG and neural oscillatory analysis, to increase the Hopfield-network storage capacity. Unlike the previously published feed-forward neural networks, our bio-inspired neural network is designed to take advantage of both biological structure and functions of the DG. To better understand the computational principles of pattern separation in the DG, we have established a mouse model of environmental enrichment. We obtained a possible computational model of the DG, associated with better pattern separation ability, by using neural oscillatory analysis. Furthermore, we have developed a new algorithm based on Hebbian learning and coupling direction of neural oscillation to train the proposed neural network. The simulation results show that our proposed network significantly expands the storage capacity of Hopfield network, and more effective pattern separation is achieved. The storage capacity rises from 0.13 for the standard Hopfield network to 0.32 using our model when the overlap in patterns is 10%.

Keywords: Hopfield network; Feedforward neural network; Pattern separation; Hippocampus DG; Neural oscillatory analysis

¹ First author.

^{*} Corresponding author. This work was supported by a grant from the National Natural Science Foundation of China (32070988)

1. Introduction

A Hopfield network is a symmetrical recurrent artificial neural network, which imitates the structure of the hippocampal CA3 region [1,2] and can be used as a device capable of storing patterns by setting the input patterns as its attractor states [3]. However, one of main limitations of a Hopfield network is its restricted storage capacity. Several approaches have been proposed to enhance the storage capacity, with some success, for example, using the absorbing-neighborhood [4] or by increasing the number of discrete states of individual Ising-Lenz neurons [5].

A recent pioneering work on a biologically plausible neural network structure has attracted our attention [6]. The work deals with modelling a similar computational function of the hippocampal dentate gyrus (DG) known as pattern separation (PS), which makes the DG produce sparse representations of spaces [7]. The basic idea of the approach is that the DG acts as a pattern separator prior to memory storage in the CA3, by which the storage capacity of a Hopfield network could be improved markedly. However, one of their biological hypotheses of 'expansion recoding' that occurs in the entorhinal cortex (EC)-DG pathway has recently been challenged because more anatomical details of DG have been revealed [8]. Furthermore, a growing number of studies have shown that complex circuits within the DG play an important role in PS [8,9,10].

Little is known about how the DG achieves PS, although several mechanisms have been proposed based on the experimental developments over the past half century. The early computational model for PS was enlightened by Marr's 'extension recoding' theory based on the fact that there is a larger population of granule cells (GCs) and a sparse firing pattern in the DG [8,11,12]. However, it is hard to elucidate the contribution of the cells with multiple place fields in the DG to PS [11]. Mossy cells (MCs) are glutamatergic cells in the DG whose somata reside in the hilus. They provide widespread feedback, namely monosynaptic excitatory and disynaptic inhibitory inputs to GCs, hence they play an important role in PS [12,13]. Chronic MC degeneration has little effect on PS [10], while contextual discrimination is impaired in the acute phase of MC degeneration. MCs show spatially selective firing fields during exploration and usually have multiple firing fields in the environment compared with GCs [8]. Meanwhile, MCs exhibit stronger place-field

remapping than GCs in response to changes in environmental cues [14]. The authors of another study that used a different method reached a similar conclusion [9]. Compared with monosynaptic exhibitory inputs to GCs, disynaptic inhibitory inputs to MCs are more essential to complete PS by GCs. In general, the complex circuit composed of a variety of neurons in the DG provides a more general explanation for the sparse activity of GCs and the mechanism of PS.

We believe that investigating how neural information propagates along the structure could lead to further understanding of the computational principles of the brain [16]. Neural oscillations regulate information flow in neural circuits and contribute to the realisation of a variety of higher brain functions, so they have been considered one of the emerging computational principles in the brain [15,17]. As the predominant oscillations in the hippocampus, theta (θ) and gamma (γ) are two of the most important rhythms because of their close association with multiple cognitive functions [18,19]. The θ and γ components of local field potentials (LFPs) in the DG are mainly associated with excitatory postsynaptic currents (EPSCs) and inhibitory postsynaptic currents (IPSCs), generated by glutamatergic neurons and interneurons (INs) respectively [20], which make it possible to infer the direction of information flow among layers in the DG.

We aimed to develop a novel feed-forward neural network inspired by the structure of the DG and neural oscillatory analysis to increase the Hopfield network storage capacity. We firstly established a mouse model of environmental enrichment (EE) and then subjected the mice to object pattern separation (OPS) and Morris water maze (MWM) tests to determine whether EE significantly alter neural activity patterns in the hippocampus. In addition, we used power spectral analysis to distinguish any change in energy distribution to better understand the computational principles of PS in the DG. Subsequently, we used two algorithms, namely PAC-CMI and PAC-PLV, to measure the strength and direction of cross-frequency coupling. Harnessing our findings on the biological structure and functions of the DG, we have proposed a bioinspired artificial neural network structure with a hidden layer representing MC in the DG. A neuron in each hidden layer has both excitatory and inhibitory projections to its input layer. For PS in the DG, we investigated the interrelationship between inhibitory and excitatory neurons.

3

The rest of the paper is organised as follows. Section 2 provides the experimental details, including experimental animals, behavioral tests, LFP recordings and neural oscillatory analysis. In addition, we discuss the anatomy of the DG neural network, and briefly introducea variation of Hebbian learning. In Section 3, inspired by coupling direction of neural oscillation, we present a novel computational model based on Hebbian learning. In Section 4, we provide the experimental results from the biological tests and our novel algorithm. Furthermore, we assess the function of the neural network and its performance including the convergence across the layers. In Section 5, we discuss the differences between our research and other related studies. Finally, we provide our conclusions in Section 6.

2. Methods and materials

2.1. Animals and treatment

It is well known that there is a positive effect of EE on PS [21,22]. In this study, male C57 mice (n = 12) at the age of weaning (21 days) were used and they were randomly divided into two groups: an enriched environment (EE) group (n = 6) and a control (CON) group (n = 6). Mice in the EE group were raised in a large, multilayer space (60 \times 40 \times 35 cm³) with 6 mice/cage. There were various toys including houses, running wheels, hammocks, scales, small bells, ladders, and tunnels available for animals in the EE group (Fig.1B, left). The toys were changed twice a week. Mice in the CON group were raised in standard cages $(36 \times 18 \times 14 \text{ cm}^3)$ with 6 mice/cage without any objects (Fig.1B right). Mice were housed in different cages until 12 weeks of age. The housing conditions, except for the cages, were kept strictly identical between these two groups. The room temperature was constant at 22 ± 2 °C. Animals had free access to food and water during the experiment. Mice were purchased from the Laboratory Animal Center, Academy of Military Medical Science of People's Liberation Army. Every procedure was followed the protocol approved by from the Committee for Animal Care at Nankai University (20160004). The schematic representation of our experimental design, which lasted 73 days, is shown in Fig.1A. Section 2.2 and 2.3 provide details about the behavioural tests.



Fig. 1. Experiments and artificial neural network. A. The schematic representation of the experimental design. B. Mice in the EE group (left panel) have larger space and various toys such as houses, running wheels. In contrast, animals in the CON group were housed in standard cages ($36 \times 18 \times 14$ cm) with 6 mice/cage. C. The experimental protocol of MWM test and OPS test. D. The anatomy layers of the dentate gyrus. E. Schematic of the DG with granule cells, mossy cells, and interneurons.

2.2. Object pattern separation

OPS was proposed by Goethem et al. [23], with the aimed to measure the hippocampus-driven PS. The test was performed in a circular arena with a depth of 40 cm and a diameter of 60 cm (Fig.1C-right). Half the circular wall is transparent, and the other half is black. The bottom of the arena contains 18 position tags. There are two tags for position 1 and four tags for positions 2 to 5. In the experiment, we performed three sessions at three positions (1, 3, and 5) with an inter-session interval of 48 h (1 session/day). Each session was divided into two trials (4 min/trial). During the first trial (T1), mice need to be familiar with the positions of the two objects located in position 1. There was no difference in appearance between the two objects. One hour later, the mice performed the second trial (T2) with one of the two objects displaced to one of three positions along the vertical axis. The red circle in Fig. 1C (right) indicates the new position of the object on the right during T2 (taking position 5 as an example). The test trial is executed at position 3, 24 h after completion of the final learning trials. The test trial process is the same as the learning trial. Finally, the time the mice stayed near the familiar object (t_f) and the new location of the object (t_n) in T2 was collected to measure the discrimination index by $(t_n - t_f)/(t_n + t_f)$.

2.3. Morris water maze test

The MWM test was executed with appropriate modifications [38]. It was performed in a circular pool with 90 cm in diameter and 60 cm in height, filled with opaque water maintained at the temperature of 20-22 °C (Fig.1C-left). The tank was divided into four equal quadrants by two imaginary vertical axes. During the initial training stage, a platform with 10 cm in diameter was located in a target quadrant, submerged 1 cm below the water surface. The animals were placed randomly into the tank facing the side wall and near the edge at one of four points randomly. The mice were allowed 60 s to find the hidden platform; if they did not find it, they were guided to the platform. After each training session, mice were allowed to stay on the platform for about 10 s. The initial training stage (IT) lasted for 5 days; both the escape latency, which is the time needed to find the platform, and swimming speed were collected. Mice that failed to find the platform were guided to the platform and stayed there for about 10 s, and the escape latency was recorded as 60 s. In the space exploration test stage (SET, day 6), the platform was removed, and the mice were placed in the pool on the opposite side of the platform. The time spent in the target quadrant and the number of crossing platform were examined in the test stage. A CCD camera connected to a dedicated computer (Ethovision 2.0, Noldus, Wageningen, Netherlands) was used to record the navigation performance during the MWM test.

2.4. Electrophysiological recording

After the behavioural tests, all the animals were anaesthetized with 30% urethane (1.2 g/kg body weight, intraperitoneal; Sigma-Aldrich, St. Louis, MO, USA), and then they were placed in a stereotaxic apparatus (Narishige, Tokya, Japan) for surgery. A bipolar stimulating electrode was slowly implanted in the angular bundle of the perforant pathway (PP: 3.0 mm lateral to the midline, 3.8 mm posterior to bregma, and 1.8 mm ventral below the dura), and the recording electrode was gently inserted into the DG (2.0 mm lateral to the midline, 1.4 mm posterior to bregma, and 1.8 mm ventral below the dura) according to the mouse brain atlas [24]. LFP signals were collected at a sampling rate of 1000 Hz for 10 min by means of a dual-

channel bioelectric amplifier and PowerLab/8SP acquisition.

2.5. LFP data analysis

2.5.1 Power spectrum analysis

Recorded LFP data were analyzed using built-in and custom-written MATLAB programs. The Multitaper Spectral Estimation in the Chronux toolbox was applied to calculate power spectra. Each LFP recording was segmented using 40-s sliding windows with 50% overlapping. The data were then Fourier-transformed with a Hamming window and averaged to obtain a direct Multitaper spectral estimate. Next, the direct estimates were normalized to power ratios of the whole frequencies (1–100 Hz) for further statistical analysis. A time-frequency power spectrogram was generated in 200-s data segments using a 20-s time-window with a 10-s overlap.

2.5.2 Phase-amplitude coupling analysis

Phase-amplitude cross-frequency coupling (PAC) intensity is the most predictive neurophysiological indicator of learning [25,26]. This measur shows the distribution of the fast oscillation amplitude within the specific phase of slow oscillations, indicating the accuracy and efficiency of integration and processing of neural information [40]. In our study, the phase locking value (PLV) of cross-frequency PAC was applied to evaluate the strength of PAC in the signal of length N and defined as:

$$PAC - PLV_{ph-phamp} = \left| \frac{1}{N} \sum_{t=1}^{N} exp\{i[\phi_{low}(t) - \phi_{high}(t)]\} \right|$$
(1)

where, N is the length of signals. The phase of the low frequency channel (4-12 Hz) was filtered by Hilbert transform and denoted as $\phi(\theta)$; then, the amplitude of the high-frequency channel (30-100 Hz) was narrow-band filtered and denoted as $\alpha(\gamma)$. The second Hilbert transform was applied to determine the phase of the highfrequency amplitude, that is, $\phi(\alpha(\gamma))$. Finally, the PLV between $\phi(\theta)$ and $\phi(\alpha(\gamma))$ was calculated. The procedure was applied on a data length of 20 s with 10-s time steps. The PAC values were normalized to obtain more robust and accurate estimates through a surrogate data approach [27,28]. Specifically, the surrogate data were generated by shuffling the high amplitude time series with 50 surrogate times and the measurement was defined as:

$$Z_score = \frac{Z - mean(Z_{surrogate})}{std(Z_{surrogate})}$$
(2)

where Z refers to the original values of the indices. The normalized indices, Z-scores, were applied in the final results. The surrogate time was set as 100.

2.5.3 Conditional mutual information

Conditional mutual information (CMI) is a measurement for detecting information flow's direction, by using the phase information derived from interacting neural rhythms recorded in related cerebral regions [29,30]. To determine the hierarchical structure of the DG, both slow and fast oscillations were extracted from the LFP signals. The phase information of both slow and fast oscillations was obtained by Hilbert transformation. ϕ_{fast} represents the phase of the envelope of the the amplitude of the filtered γ frequency signal and ϕ_{slow} denotes the phase of the filtered θ frequency signal. The CMI could be yielded from:

$$I(\phi_{slow}; \Delta\phi_{fast} | \phi_{fast})$$

= $H(\phi_{slow} | \phi_{fast}) + H(\Delta\phi_{fast} | \phi_{fast})$
 $- H(\phi_{slow}, \Delta\phi_{fast} | \phi_{fast})$ (3)

$$I(\phi_{fast}; \Delta \phi_{slow} | \phi_{slow})$$

= $H(\phi_{fast} | \phi_{slow}) + H(\Delta \phi_{slow} | \phi_{slow})$
 $- H(\phi_{fast}, \Delta \phi_{slow} | \phi_{slow})$ (4)

where $H(\cdot)$ is defined as defined by Shannon, and $\Delta \phi = \phi(t + \tau) - \phi(t)$. Each LFP recording was segmented using 24-s sliding windows with 50% overlapping. The phase increment was set to $\tau = 100 \text{ ms}$. The normalized directionality is defined as:

$$D_{fast \to slow} = \frac{A - B}{A + B},$$
(5)
where $A = I(\phi_{fast}; \Delta \phi_{slow} | \phi_{slow})$
 $B = I(\phi_{slow}; \Delta \phi_{fast} | \phi_{fast})$

2.6. Anatomy of the dentate gyrus

The schematic interconnection between the EC, the DG and CA3 has been discussed in several studies[31,32,33]

and is shown in Fig. 1D and E. It includes a perforant path (PP) projection from the second layer of the EC to the DG, and CA3, a mossy fibre (MF) projection from the DG to CA3, and a backward loop from CA3 to the hilus of the DG. Recent evidences has revealed that MCs and several types of inhibitory neurons in the hilus of the DG may play an important role in achieving PS [12,13,34]. An MC receives the local GC input (Fig.1D) [9,33,35], and its activation leads to the inhibition of the local GC by inhibitory neurons and inform non-local GCs by recurrent collaterals [33,36]. MC degeneration causes transient DG hyperexcitability and impaired PS [10]. Moreover, disynaptic inhibition is more important than monosynaptic excitation to implement PS [9]. A similar finding has also been reported: weak inhibitory neurons mediated by MCs are the culprit of temporal lobe epilepsy (TLE) [14]. However, it is still unclear how MCs are involved in PS because this process is dynamic and not necessarily dependent on the anatomy of connection.

2.7. Expanded Hebbian learning

Hebbian learning has been proposed to describe the principles of learning at the synaptic level [37]. Recently, a novel expansion of the Hebbian learning method, named expanded Oja's rule, has been proposed [41]. It employs the merits of BCM theory simulating long-term depression (LTD) [39] and completely solves the issue that the weight increases without boundary. The learning rule is illustrated as below:

$$\pi_L \frac{d\boldsymbol{W}_i}{dt} = g(Q)(R^p \boldsymbol{v}_i - \langle \boldsymbol{W}, \boldsymbol{v} \rangle W_i)$$
(6)

where
$$Q = \frac{\langle W, v \rangle}{\langle W, W \rangle^{\frac{p-1}{p}}}$$
 and $g(h) = \begin{cases} 0, & h < 0 \\ -\Delta, & 0 \le h < h, \\ 1, & h_* \le h \end{cases}$

In this equation, g(h) is the activation function of any real number input h. The constant τ_L is defined as the time scale of learning dynamics. The parameter Wrepresents the weight matrix. Output of the network Q is formed by the input vector v transformed by W and divided by the p-norm of W. The above procedure shows that each row of W finally converges to a sphere of radius R. When Q is less than threshold h_* , g(h) inhibits Wwith the extent ' $-\Delta$ ', otherwise, it excites W. In this study, we have proposed a new learning algorithm based on Oja's rule.

2.8. Statistics

All analyses were performed by using SPSS statistics 21.0 (IBM Corp., Armonk, NK, USA) with the significance level set at 0.05. All data are presented as the mean \pm S.E.M. A one-way analysis of variance (ANOVA) or a two-way repeated ANOVA was employed, followed by Bonferroni's multiple-comparison test to determine between-group differences. The Kruskal-Wallis test was employed for data that did not follow a Gaussian distributions (P \leq 0.05),.

3. Proposed works

3.1. Enhancement of pattern separation associated with information flow

To evaluate the effect of EE on PS, the OPS test was performed on mice in the EE group and the CON group (see Section 2.2). The mice in the CON group were trained in a circular arena until a reliable OPS performance was established (Fig.2A). The discrimination index in the EE group was greater than that in the CON group, suggesting that the EE mice have a stronger PS ability compared with the CON mice (Fig.2B).

The MWM test was used to assess the hippocampusdependent spatial learning and memory ability [42,43]. In the IT stage, animals were subjected to 5 consecutive days of training to learn the location of a hidden platform (Fig.2C). A two-way repeated ANOVA showed significant housing environment [F (1,10) = 9.33, P < 0.01] and day [F (4,54) =38.55, P < 0.001] effects, but the the housing environment × day interaction was no significant [F (4,54) = 1.14, P > 0.05]. There were lower latencies on days 3-5 (Fig.2C, P < 0.05) in the EE group compared with the CON group. In the SET, there were more platform crosses in the EE group compared with the CON group (Fig.2D, P < 0.05).

LFPs were collected from the DG in the EE and CON groups. To investigate how EE impacted the brain rhythms of LFPs, the power spectrum density (PSD) was measured. Fig. 2E shows the averaged PSD in both the CON and EE groups: EE clearly altered the PSD distribution. There was a significant increase in either the θ frequency band (Fig. 2F-left, P < 0.05) or high- γ rhythm (Fig.2F, right, P < 0.05) in the DG. To better characterise the EE-mediated change in the PSD distribution, the growth rate of power was defined as: [(band-power in EE) / (band-power in CON)].

Interestingly, the increase in power was greater at the high- γ band than that at the θ band (Fig.2G, P < 0.05).

PAC intensity could be used to characterise an interaction between two different rhythms. θ - γ PAC is an integration mechanism for different spatial and temporal scales in the cognitive process. The strength of PAC in the DG was higher in the EE group than that in the CON group (Fig.2H-left, P < 0.05), indicating that EE significantly augmented neural information integration and processing. Meanwhile, there was no significant change between these two groups in the PP region (Fig. 2H-right).

Furthermore, CMI was applied to infer the direction of cross-frequency coupling in the DG (Fig.2I). The inset in Fig. 2I shows that there was greater directional coupling between the high- γ rhythm and the θ rhythm in the EE group than that in the CON group. The data illustrates that the activity of low-frequency signals is more capably regulated by the high-frequency signals in the EE group compared with the CON group. This finding suggests that there is a PS function in the DG. However, the above-mentioned regulatory capacity was not observed in the PP region (Fig. 2J), implying that the PP does not function in PS.

3.2. A novel computational model based on the dentate gyrus structure and function

We have proposed a novel dentate gyrus computational model (DGCM) based upon the structure and function of the DG, including a network structure inspired by the real DG and an algorithm utilising Hebbian learning. The biological connections in the DG, including GCs, MCs, and Ins, are considered as the basic structure of the artificial neural network (ANN, Fig. 1E). The structure of our proposed network is shown in Fig. 3A. To better illustrate the processing order of the network to input data, the network is separated into three layers (Fig. 3B). The first layer is called the input-output (I/O) layer, which represents GCs, while the second layer is named as the hidden layer, which mixes the biological functions of MCs and INs. The learning process is discussed next.

3.2.1 Learning process from the input-output layer to the hidden layer (GC \rightarrow MC)

The hidden layer receives the data from the I/O layer to extracting the most representative features of the local



Fig. 2. **Results of biological experiments. A.** Means + SEM for the d2 index in the OPS on 1, 3, 5 positions for the pilot study with control C57/BL mice. **B.** Means + SEM for the d2 index in the OPS for EE treatment on position 3. **C.** Mean+ SEM for escape latency was determined for the IT stage in MWM test among two groups. **D.** Mean+ SEM for number of platform crossings in the SET stage of MWM test. **E.** Averaged power spectrum within groups from 1 Hz to 100 Hz in DG region. **F.** The histogram of PSD at θ (4-12 Hz, left panel) and High γ (50-90 Hz, right panel). **G.** Ratio computed by averaged band power of CON divided by band power of EE. **H.** Histogram of Mean + SEM for PAC-PLV of DG and PP. **I.** Line chart with Mean + SEM which indicate the PAC_CMI between high γ and θ by within DG. The embedded figure illustrates the averaged PAC_CMI of high γ to θ . **J.** Line chart with Mean + SEM which indicate the PAC_CMI of high γ to θ .

7



Fig. 3. The architecture of DGCM and its flowchart. A. The architecture of DGCM. B. A network structure unfold diagram. C. Stepby-step DGCM work through: *i*. The GCs receive input patterns from the entrochinal cortex; *ii*. The pattern of GCs actives MCs and INs; *iii*. The MCs uses INs to realize *k*-WTA and renew the synaptic w_{fwd} between GCs and MCs; *iv*. The INs communicate local inhibitory information to GCs and updated inhibitory synaptic weight w_{inh} ; *v*. Updating the excitatory synaptic weight w_{exc} by combining inhibitory weight and GCs activity; *vi*. The output of GCs are regulated by MCs.

input, as shown in Fig. 3C(i - ii). Note that w_{fwd} , from the I/O layer to the hidden layer, is not fully connected. Advances in neurobiology have shown that MCs usually have several place fields [8]. Around 47% of MCs are recruited into the active neuronal network at the same time during sharp wave ripples (SWRs) [34], implying that soft competitive learning may better simulate the activity of MCs than hard competitive learning. As a result, we make use of the activation function of 'Neural-Gas' [45] and the framework of k-WTA [31] as our soft competition learning scheme for w_{fwd} . The updated rule for w_{fwd} is defined as follows:

$$\frac{dw_{fwd}^{i}}{dt} = a \cdot e^{-\frac{k_{i}}{\lambda}} \cdot (G_{i} - w_{fwd}^{i})$$
(7)

Note that the learning ratio $a \in [0,1]$ describes the extent of the modification and $exp(-k_i/\lambda) \in [0,1]$ denotes post-synaptic strength, where k_i and λ represent arrangement and the number of neural units, respectively. G_i is the local input vector received by the *i*th hidden neuron, which has the weight vector w_i . The

closer G_i is to w_i , the higher the sort of k (close to zero) and the greater the postsynaptic membrane response. The larger the dw/dt is, the closer the updated w_i and G_i will be. Therefore, each neuron obeying the update rules of Eq. (7) can converge to a stable value that describes the local feature of the input vector.

When the hidden layer is activated, only the first K neurons that are most active are retained, while the other hidden layer neurons are inhibited by INs through the k-WTA mechanism. Then, the network updates the w_{fwd} according to Eq. (7) with GCs as the presynaptic membrane and MCs as the postsynaptic membranes (Fig. 3C(iii)).

3.2.2 Learning process from the hidden layer to the input-output layer ($MC \rightarrow GC$)

Whenever the neuron in the hidden layer is activated, a bilateral impact including both excitation w_{exc} and inhibition w_{inh} will pass back to the I/O layer through disynaptic inhibition and monosynaptic excitation, as shown in Fig. 3B. The only issue is which of them should

be updated first. The higher PAC-CMI value in the EE group suggests that inhibitory neurons have a better regulatory ability than excitatory neurons. The main reasons are: 1) as shown in Eqs (3) and (4), the directionality of information flow comes from a timedelay system; and 2) in the DG, fast oscillation is mostly derived from the activity of inhibitory neurons, while the slow oscillation is primarily derived from the activity of excitatory neurons [20]. Therefore, w_{inh} is updated first (as shown in Fig. 3C(iv)), which is then used to update w_{exc} (illustrated in Fig. 3C(v)).

The updated rules for inhibition synapse can be defined as follows:

$$\frac{dw_{inh}^{lk}}{dt} = -\frac{\alpha\rho(w_{exc}, b)}{\left\|\left(M, M^{l_k}\right)\right\|^2} \left(M^{l_k} + w_{inh}^{l_k}\rho(w_{exc}, b)\right)$$
where
$$(8)$$

$$\rho(w_{exc}, b) = (w_{exc} \quad b) \begin{pmatrix} M \\ \|(M, M^{l_k})\| \end{pmatrix}$$

The updated rules for excitation synapse are:

$$\frac{d(w_{exc}, b)}{dt} = \frac{\alpha \sigma(w_{exc}, b)}{\|(M, M^{l_k})\|^2} \begin{bmatrix} M \\ 1 \end{bmatrix}$$
$$-(w_{exc}, b)\sigma(w_{exc}, b) \end{bmatrix} \qquad (9)$$

where

 $\sigma(w_{exc}, b) = ReLu(\rho(w_{exc}, b) + M^{l_k} w_{inh}^{l_k})$

Note that $w_{inh}^{l_k}$ is the weight of the inhibitory vector with local index l_k , while w_{exc} is the weight of the excitatory vector. They correspond separately to the local input M^{l_k} and M, and finally converge to the frequencyinput pattern. b is used to model the threshold of DC and converges to the average fire rate of I/O neurons. Similarly to Eq. (7), α represents the learning rate.

When the hidden layer neurons are activated, they deliver inhibitory effects to the I/O layer firstly and update w_{inh} according to Eq. (8) based on the response of the postsynaptic membrane (GCs' of Fig. 3C(iv)). Note that the inhibitory synapses of the hidden layer to the I/O layer are also not fully connected. This is because the range of influence of GABAergic interneurons (INs) is usually local [50]. GCs with inhibitory components are used to update the excitatory synapses from the hidden layer to the I/O layer (Fig. 3C(v)). Unlike the inhibitory synapses, the excitatory synapses of the hidden layer to the I/O layer are

fully connected. This is because MCs make synaptic connections with around 75% of GCs [33]. Therefore, it is more appropriate to simulate this phenomenon by the dense connections. Finally, influenced by both excitation and inhibition, GCs maintain an excitatory output (Fig. 3C(vi)). The description of the algorithm for PS is illustrated as Table 1.

Table 1. Algorithm for pattern separation

Algorithm 1: Pattern separation
Hyperparameters:
learning rate: α
input /output layer size:n
window length: l_w
overlapping rate:r
number of input patterns: q
input pattern: $P = \{\vec{p}_1,, \vec{p}_k,, \vec{p}_q\}$
input vector: $\vec{p}_k = (x_1, x_2, \dots, x_n)$
<code>'repeat'until threshold $heta$ has been satisfied</code>
$W_{GC \to MC} = \{\vec{h}_1, \dots, \vec{h}_k, \dots, \vec{h}_m\},\$ where $\vec{h}_{k} = (w^{fwd} w^{fwd} - w^{fwd})^T$
$w_{leven} = (w_1, w_2,, w_{l_w})$
$W^+_{MC \to GC} = \{ \vec{g}^+_1, \dots, \vec{g}^+_k, \dots, \vec{g}^+_n \},\$
where $\vec{g}_{k}^{+} = (w_{1}^{+}, w_{2}^{+},, w_{m}^{+})^{T}, m = \left \frac{n - l_{w}}{l_{w} - r}\right + 1$
$W^{MC\to GC} = \{\vec{g}^1, \dots, \vec{g}^k, \dots, \vec{g}^n\},\$
where $\vec{g}_k^- = (w_1^-, w_2^-, \dots, w_s^-)^T$, $s = \left\lfloor \frac{n - r \cdot l_w}{l_w - r \cdot l_w} \right\rfloor$
$\vec{b} = (b_1, b_2, \dots, b_n)$
START
Initialization{
for each $W \in \{W_{GC \to MC}, W^+_{MC \to GC}, W^{MC \to GC}, \vec{b}\}$ do $W = random\{0 \sim 1\}$
}
foreach $\vec{p}_k \epsilon P$ do
learning rules for $GC \rightarrow MC$
$\vec{p}_k = (x_1, x_2, \dots, x_n)$
repeat
for <i>i</i> th hidden neuron weight $ec{h_i}$
$j = (i - 1)(l_w - r) + 1,$
$k_i = sort\{\vec{h}_i \cdot \vec{p}_k[j:j+l_w-1]\}$
$\frac{d\vec{h}_i}{dt} = a \cdot e^{-\frac{k_i}{m}} \cdot (\vec{p}_k[j:j+l_w-1] - \vec{h}_i) \qquad \text{eq. (7)}$
update \vec{h}_i as $\vec{h}_i + \frac{d\vec{h}_i}{dt}$
end
until $ W_{GC \to MC} _2 < \theta$
learning rules for MC $\xrightarrow{+}$ GC and MC $\xrightarrow{-}$ GC
$\vec{M} = W_{GC \to MC} \cdot \vec{p}_k$
repeat
for L_{ν} th inhibitory input weight a_{1}^{-}

$$\rho(W_{MC \to GC}^{+}, \vec{b}) = (W_{MC \to GC}^{+}, \vec{b}) \begin{pmatrix} M \\ \|(\vec{M}, \vec{M}^{l_{k}})\| \end{pmatrix}$$
$$\frac{dg_{\vec{l}_{k}}}{dt} = -\frac{\alpha \rho(W_{MC \to GC}^{+}, \vec{b})}{\|(\vec{M}, \vec{M}^{l_{k}})\|^{2}} (\vec{M}^{l_{k}} + \vec{g}_{\vec{l}_{k}} \rho(W_{MC \to GC}^{+}, \vec{b}))$$

eq. (8)

eq. (9)

end

for \boldsymbol{k} th excitatory input weight g_k^+

$$\sigma(\vec{g}_k^+, \vec{b}) = ReLu\left(\rho(\vec{g}_k^+, \vec{b}) + \vec{M}^{l_k} \vec{g}_{l_k}\right)$$

$$\frac{d(\vec{g}_k^+,\vec{b})}{dt} = \frac{\alpha\sigma(\vec{g}_k^+,\vec{b})}{\|(\vec{M},\vec{M}^{l_k})\|^2} \left[(\vec{M} - 1) - (\vec{g}_k^+ - \vec{b})\sigma(\vec{g}_k^+,\vec{b}) \right]$$

end

update $(W_{MC \to GC}^+, \vec{b})$ as $(W_{MC \to GC}^+, \vec{b}) + \frac{d(\vec{g}_k^+, \vec{b})}{dt}$ update $W_{MC \to GC}^-$ as $W_{MC \to GC}^- + \frac{d\vec{g}_{l_k}^-}{dt}$ until $\|W_{MC \to GC}^+\|_2 < \theta$ & $\|W_{MC \to GC}^-\|_2 < \theta$ end Return $\{W_{GC \to MC}, W_{MC \to GC}^-, W_{MC \to GC}^+, \vec{b}\}$

According to each pattern $p_k \epsilon P$, learning rules for GC \rightarrow MC aim to generate the hidden layer state vector M. Firstly, the PP input activates its corresponding neurons, in the local hidden layer. According to Eq. (7), the more intense the activation of the hidden layer neuron, the more adequate the update of its synaptic weight, meaning that the weight is closer to the input. Predictably, when hidden layer neurons are activated, their corresponding local input charateristics appear more frequently. When the weight of $GC \rightarrow MC$ stabilises, the training of $MC \rightarrow GC$ begins. Second, the activity state of hidden layer neurons has been taken as input. The state of the original input layer is changed by the activation of the hidden layer, which can be divided into the excitatory and the inhibitory, as defined in Eqs (8) and (9). If the GC activity is greater than the threshold, the neurons remain active after the dual influence of excitation and inhibition are superimposed on the GCs. Otherwise, GCs activity is suppressed.

We have proposed the novel DGCM to increase the storage capacity of a Hopfield network. Moreover, we have used the same input data to compare between DGCM and the Hopfield network.



Fig. 4. Results of ANN test. A. Loss of feedforward weight from the first layer to the second layer. B. Loss of feedback excitatory weight from the second layer to the first layer. C. Loss of feedback inhibitory weight from the second layer to the first layer. D. Loss of bias of the first layer. E. The histogram of storage capacity of three models under the condition of pattern overlapping percentage equals to 10. F. The histogram of storage capacity of three models under the condition of pattern overlapping percentage equals to 40. G. Number of storage patterns of standard Hopfield network as the function of input noise in patterns. H. Number of storage patterns of DGCM+Hopfield network as the function of input noise in patterns. I. Number of storage patterns of standard Hopfield network as the function of patterns complexity. J. Number of storage patterns of DGCM+Hopfield network as the function of patterns complexity.

4. Results

4.1. The dentate gyrus computational model enhances the Hopfield network storage capacity

We investigated whether the Hopfield network storage capacity is enlarged marketly by using our new modelling algorithm. Input data are generated randomly for q different patterns. Each pattern is a 100-bit binary string. q patterns are orthogonal to each other, and each pattern has 100/q bits as 1 and the remaining bits are 0. The dot product of each pair of the pattern is 0. q is chosen to be bigger than 1 and less than 40, and its range is fully orthogonal with 50% overlapping. The overlapping rate is achieved by adding random variation to the completely orthogonal mode. Because q modes generated each time have randomness, each q mode is stimulated 1000 times to reduce the impact of randomness on our network performance evaluation. When the network can fully recall q stored patterns, a counter for the correct number of recalls is increased by 1; otherwise, it remains unchanged. Finally, the recall accuracy is obtained as the counter divided by 1000. The parameter values are summarised in Table 2.

Table 2. Parameters

Variables	Value		
α	0.1		
п	100		
l_w	16		
r	0.5		
q	1~40		
θ	0.01		

The learning rate α is set to 0.1, commonly used in other ANN models. *n* is set to 100, referred to Ref. 6. The maximum number of *q* is 40 since the maximum storage capacity is 0.32. When θ equals 0.01, the model can achieve good performance, suggesting that it is not necessary to set it as smaller value. l_w and *r* are adjusted according to our experience.

The learning curves representing loss over time show the performance of the model: the lower the loss, the better the model performance. The learning curve of w_{fwd} corresponding to the left-hand side of Eq. (7) is presented in Fig. 4A. Furthermore, the learning curve of w_{inh} corresponding to the left-hand side of Eq. (8) is shown in Fig. 4B. In addition, the learning curves of w_{exc} and **b** corresponding to the left-hand side of Eq. (9) are given in Fig. 4C and 4D, respectively. Fundamentally, the loss evidently decreases over time in all four curves; therefore, the model is capable of learning. The results of convergence test are shown in Fig. 4A-D. Note that the convergence of $W_{GC \to MC}$ (Fig. 4A) corresponds to the left side of Eq. (7). The convergence of $W_{\overline{MC} \to GC}$ (Fig. 4B) corresponds to the left side of Eq. (8). The convergence of $W_{MC \to GC}^+$ and **b** (Fig. 4C and D, respectively) corresponds to w_{exc} and **b** in Eq. (9).

Three models, including a standard Hopfield network (Fig. 4E and F, left), a previous bio-inspired model [6] (Fig. 4E and F, middle) and our new model (Fig. 4E and F, right), are considered for comparison. Fig. 4E and F display the number of storage patterns in these three models under the condition of pattern overlapping of 10% and 40%. Note that the overlap in patterns has been calculated by using the Hamming distance. Compared with the previous bio-inspired model, our new model significantly increased the number of storage patterns from 18 to 32 when there is 10% overlap (Fig. 4E). Our new model still slightly enhanced the storage capacity when there is 40% overlap (Fig. 4F). Additional details of comparison among these three models are shown in Table 3.

Table 3. Parameters

OP	10	20	30	40	50
Our model	32	22	16	11	9
Ref. [6]	18	~ 16	~ 14	10	N.G.
Std. Hop. Net.	13	8	6	3	0

"~" represents the estimated value from Ref.[6], while N.G. represents "no result given". The storage capacity of all three models is decreased with the continuous increase of OP. When the OP is 40, the similar performance has been obtained by both our model and Ref.[6]. When the OP is 50, our model still achieves the storage capacity compared to other two models.

The number of storage patterns as a function of the overlap percentage in a standard Hopfield network and a Hopfield network with the novel DGCM is represented in Fig. 4G and Fig. 4H, respectively. The background color represents the success rate, which is the number of storage modes retrieved correctly over the total number of retrievals. A success rate of 1 indicates all the retrieved patterns were correct, while a success rate of 0 indicates all the retrieved patterns were incorrect. The curves, represented in Fig. 4G and H, correspond to the contour



Fig. 5. Quantitative evaluation of DGCM performance. A. The loss of excitatory and inhibitory synaptic weights in two orders. B. The neuronal activity of input/output layer before DGCM processing. C. The sparsity and similarity between input and output patterns. D. The neuronal activity of input/output layer after DGCM processing.

line of a 95% success rate. We measured the number of storage capacity when the overlap percentage was 10%. The storage capacity was 0.13N for a standard Hopfield network, however, it was evidently increased to 0.32N for a Hopfield network with the novel DGCM (Fig. 4H). In addition, we compared a standard Hopfield network (Fig. 4I) and a Hopfield network with the novel DGCM (Fig. 4J)

in term of the storage capacity, considering the pattern complexity as the variation. Pattern complexity is defined as the ratio of the number of bits that equal to 1 to the total number of bits in the pattern. The background color and the curves have the same meaning as in Fig. 4G and H. The results clearly show that the storage capacity of the network is reduced as the pattern complexity increases.

12

However, compared with the standard Hopfield network, the storage capacity of the new network could be effectively enhanced by the DGCM by suppressing the excitatory neurons and increasing orthogonalization among patterns.

4.2 Evaluation of the dentate gyrus computational model

We next carefully evaluated the DGCM performance. To show that the guiding role of inhibitory neurons to excitatory neurons in the DGCM is necessary for the learning process of this network, we switched the activation sequence of the two types of neurons in the learning process and retested it on the dataset. Fig. 5A shows the loss of excitatory and inhibitory synaptic weights in two different orders. The change in the activation sequence of these two types of neurons had a significant, negative impact on the convergence of inhibitory neurons (Fig. 5A, upper panel, Kruskal-Wallis test, P < 0.05), but it had almost no effect on the learning process of excitatory neurons. Furthermore, Fig. 5B and D show both the neuronal activity pattern of the I/O layer without hidden layer regulation and the pattern of the I/O layer after hidden layer regulation. There was sparser neuronal activity pattern of the I/O layer regulated by the hidden layer, which was quantitatively evaluated (Fig. 5C, left, Kruskal-Wallis test, P < 0.001). After the regulation of the DGCM hidden layer, the network activity was significantly sparser and there was less similarity among patterns (Fig. 5C, right, Kruskal–Wallis test, P < 0.001), suggesting that the reduced similarity between the two patterns is the computational biological definition of PS. Therefore, the DGCM functionally implements PS.

5. Discussion

Brain-inspired neural networks have been studied extensively in recent years due to the advancements in neuroscience [46,50]. In this study, we established a mouse model of EE to verify its effectiveness in enhancing cognitive ability. The EE group showed superior PS compared with the CON group. Meanwhile, in the DG there was greater γ to θ information flow in the EE group than in the CON group. This suggests that the neural information flow from γ to θ is associated with PS. Accordingly, we developed a new feedforward artificial neural network and found that it enhanced the storage capability of a Hopfield networks. Clearly, the novel algorithm has been inspired by the coupling direction of neuronal oscillation and has been developed by using Hebbian learning, which is different from other multilayer feedforward learning algorithms[48,49].

5.1. Comparison between a Hopfield network with the dentate gyrus computation model and other models

We developed a Hopfield network with the novel DGCM based on the study by Janarthanam et al. [6]. Those authors tried to use a DG-inspired network structure to increase the storage capacity of a Hopfield network. One of their key hypotheses was 'expansion recoding', which is derived from Marr's hypothesis of the cerebellar input layer, developed in the 1960s. A remarkable feature of the input layer is the extensive divergence from a smaller number of MF inputs to a much larger number of GCs. Because there are comparable characteristics of a pathway from the EC to the DG, we extended the above hypothesis to the DG, which is capable of PS. There are several different approaches to enhance the capacity of a Hopfield network. Kobayashi et al. improved the capacity of a Hopfield network by increasing the discrete state of each neuron [5]. Another way to enhance the storage capacity is by using non-monotonic continuous neurons, although this approach increases the computational and storage burden [2]. Another approach to increase the capacity of a Hopfield network is called absorbing neighbourhood [4]. Moreover, a computational model was developed based on the anatomical structure of the DG [44,47], which was used to infer how PS was executed. The required parameters of the model were set based on the currently known information about the real DG rather than dynamic learning. In their subsequent study, they focussed on the effect of CA3 on the hilus of the DG. CA3 improved both PS and storage capacity; therefore, the model was a better fit to empirical data. Compared with other models [31,32,51,52], the main advantage of our model is that the synaptic parameters are determined by the differential equation inspired by directional cross-frequency coupling rather than the constants given by the random number generator. Moreover, the establishment of the new model depends not only on its corresponding anatomical structure, but also on the pattern of neural oscillations, which can signify neuronal synaptic plasticity.

Another work relevant to our study was carried out by O'Reilly *et al.* [31]. They were the first researchers to

propose the k-WTA algorithm, which had been used to reveal that CA1 contains neurons with several place fields. Compared with the traditional competitive learning methods, their computational model allows k neurons to be activated in the second layer of the neural network. This leads us to believe that a soft competition algorithm could be used to simulate a special competitive environment, where there are multiple neurons that could be activated in the second layer of the network.

5.2. The role of inhibitory neurons in the learning process and excitatory-inhibitory assemblies

Recently, there has been increasing attention devoted towards understanding the role of inhibitory neurons in the learning process. In optimised artificial neural networks, each neuron makes both excitatory and inhibitory connections with its targets separately. In physiological neurons, however, every synaptic connection is either excitatory or inhibitory [53]. Several studies have attempted to explain the role of inhibitory neurons in these excitatory-inhibitory (E-I) assemblies. We have shown that inhibitory neurons modulate the activity of excitatory neurons in the mouse DG. Accordingly, excitatory neurons are able to segment similar patterns in a stimulus-specific manner. Interestingly, similar results can also be found in thalamus [54] and the primary visual cortex (VI) [55] in mice. The GABAergic synapses of basal ganglia structure area X can precisely drive the spike of neurons in the dorsolateral thalamic nucleus (DLM) [54]. Similarly, Owen et al. not only talked about the contribution of PV intermediate neurons to shaping V1 computation, but also argued that the mechanism behind this phenomenon is E-I balance [55].

The two requirements [56] to achieve E-I balance are: 1) neurons should be connected sparsely and randomly, and 2) the strength of inhibitory connections should be high enough to balance the network, in which feedback excitation and inhibition could be cancelled. The underlying mechanism regarding the unbalanced distribution between excitatory neurons and inhibitory neurons is the update rule of Eq. (9) showing that there are a small number of inhibitory synapses with a stronger connection strength to achieve E-I balance. However, the physiological significance of E-I balance to achieve PS is still not fully understood. A computational model [57] showed that the balance between excitation and inhibition was required for high-capacity, and noise-robust neuronal selectivity. Moreover, the latest evidence suggests that GC activity could enhance the precision of memory recall [58]. Taken together, the E-I balance may be an important

mechanism to achieve accurate coding in the DG.

When the information is not distributed uniformly in the feature space, discovering the effective information hidden in the feature space could greatly improve the performance of machine learning [59]. This makes it possible to improve learning performance by using local information and non-homogeneity in the training sample, such as dividing the original problem into smaller parts [60]. The probabilistic neural network was improved by using local decision circles [61]. When the information is not distributed uniformly in the feature space, discovering the effective information hidden in the feature space can greatly improve the performance of machine learning [59]. However, the hidden layer structure has a non-adaptive window length and overlap ratio. Therefore, it is necessary to establish a dynamic approach to determine the topology of hidden layer integration[62].

6. Conclusion

Advances in neuroscience offer artificial intelligence researchers an unique perspective. We hve proposed a new algorithm to enhance the storage capacity of a Hopfield network. Input layer neurons, representing cross-pattern local features, have more chances to activate excitatory hidden layer neurons and to be inhibited by their downstream inhibitory neurons. This preserves exclusive local features for each mode and brings sparser and more orthogonal inputs to a Hopfield network. Potential improvement directions include the introduction of a probabilistic neural network or a dynamic ensemble learning algorithm.

References

- de Almeida, L., Idiart, M. and Lisman, J.E., 2007. Memory retrieval time and memory capacity of the CA3 network: role of gamma frequency oscillations. Learning & memory, 14(11), pp.795-806.
- Storkey, A., 1997. Increasing the capacity of a Hopfield network without sacrificing functionality. Proceedings of the International Conference on Artificial Neural Networks, pp.451-456.
- Kube, K., Herzog, A. and Michaelis, B., 2006, September. Increased storage capacity in Hopfield networks by smallworld topology. In International Conference on Natural Computation(pp. 111-114). Springer, Berlin, Heidelberg.

- Gosti, G., Folli, V., Leonetti, M. and Ruocco, G., 2019. Beyond the maximum storage capacity limit in Hopfield recurrent neural networks. Entropy, 21(8), p.726.
- Kobayashi, M., 2019. Storage capacity of hyperbolic Hopfield neural networks. Neurocomputing, 369, pp.185-190.
- Janarthanam, V.V., Vishwanath, S. and Shanthi, A.P., 2020. A biologically plausible network model for pattern storage and recall inspired by Dentate Gyrus. Neural Computing and Applications, 32(17), pp.13289-13299.
- Kumaran, D., Hassabis, D. and McClelland, J.L., 2016. What learning systems do intelligent agents need? Complementary learning systems theory updated. Trends in cognitive sciences, 20(7), pp.512-534.
- GoodSmith, D., Chen, X., Wang, C., Kim, S.H., Song, H., Burgalossi, A., Christian, K.M. and Knierim, J.J., 2017. Spatial representations of granule cells and mossy cells of the dentate gyrus. Neuron, 93(3), pp.677-690.
- Danielson, N.B., Turi, G.F., Ladow, M., Chavlis, S., Petrantonakis, P.C., Poirazi, P. and Losonczy, A., 2017. In vivo imaging of dentate gyrus mossy cells in behaving mice. Neuron, 93(3), pp.552-559.
- Jinde, S., Zsiros, V., Jiang, Z., Nakao, K., Pickel, J., Kohno, K., Belforte, J.E. and Nakazawa, K., 2012. Hilar mossy cell degeneration causes transient dentate granule cell hyperexcitability and impaired pattern separation. Neuron, 76(6), pp.1189-1200.
- Cayco-Gajic, N.A. and Silver, R.A., 2019. Re-evaluating circuit mechanisms underlying pattern separation. Neuron, 101(4), pp.584-602.
- 12. Nakazawa, K., 2017. Dentate mossy cell and pattern separation. Neuron, 93(3), pp.465-467.
- Scharfman, H.E., 2016. The enigmatic mossy cell of the dentate gyrus. Nature Reviews Neuroscience, 17(9), pp.562-575.
- Bui, A.D., Nguyen, T.M., Limouse, C., Kim, H.K., Szabo, G.G., Felong, S., Maroso, M. and Soltesz, I., 2018. Dentate gyrus mossy cells control spontaneous convulsive seizures and spatial memory. Science, 359(6377), pp.787-790.
- Sejnowski, T.J. and Paulsen, O., 2006. Network oscillations: emerging computational principles. Journal of Neuroscience, 26(6), pp.1673-1676.
- Lynn, C.W. and Bassett, D.S., 2019. The physics of brain network structure, function and control. Nature Reviews Physics, 1(5), pp.318-332.
- Li, Q., Yang, C., Zhang, X., Yang, Z. and Zhang, T., 2019. Arginine vasopressin attenuates dysfunction of hippocampal theta and gamma oscillations in chronic cerebral hypoperfusion via V1a receptor. Brain research bulletin, 153, pp.84-92.
- Zheng, C. and Zhang, T., 2015. Synaptic plasticity-related neural oscillations on hippocampus-prefrontal cortex pathway in depression. Neuroscience, 292, pp.170-180.
- Nishida, H., Takahashi, M. and Lauwereyns, J., 2014. Within-session dynamics of theta–gamma coupling and high-frequency oscillations during spatial alternation in rat

hippocampal area CA1. Cognitive neurodynamics, 8(5), pp.363-372.

- Pernía-Andrade, A.J. and Jonas, P., 2014. Theta-gammamodulated synaptic currents in hippocampal granule cells in vivo define a mechanism for network oscillations. Neuron, 81(1), pp.140-152.
- Clemenson, G.D., Henningfield, C.M. and Stark, C.E., 2019. Improving hippocampal memory through the experience of a rich Minecraft environment. Frontiers in behavioral neuroscience, 13, p.57.
- Bekinschtein, P., Oomen, C.A., Saksida, L.M. and Bussey, T.J., 2011, July. Effects of environmental enrichment and voluntary exercise on neurogenesis, learning and memory, and pattern separation: BDNF as a critical variable?. In Seminars in cell & developmental biology (Vol. 22, No. 5, pp. 536-542). Academic Press.
- van Goethem, N.P., van Hagen, B.T. and Prickaerts, J., 2018. Assessing spatial pattern separation in rodents using the object pattern separation task. Nature protocols, 13(8), pp.1763-1792.
- 24. Paxinos, G. and Watson, C., 2006. The rat brain in stereotaxic coordinates: hard cover edition. Elsevier.
- Burke, S.N., Wallace, J.L., Nematollahi, S., Uprety, A.R. and Barnes, C.A., 2010. Pattern separation deficits may contribute to age-associated recognition impairments. Behavioral neuroscience, 124(5), p.559.
- Bolz, L., Heigele, S. and Bischofberger, J., 2015. Running improves pattern separation during novel object recognition. Brain Plasticity, 1(1), pp.129-141.
- Canolty, R.T., Edwards, E., Dalal, S.S., Soltani, M., Nagarajan, S.S., Kirsch, H.E., Berger, M.S., Barbaro, N.M. and Knight, R.T., 2006. High gamma power is phase-locked to theta oscillations in human neocortex. Science, 313(5793), pp.1626-1628.
- Bruns, A. and Eckhorn, R., 2004. Task-related coupling from high-to low-frequency signals among visual cortical areas in human subdural recordings. International journal of psychophysiology, 51(2), pp.97-116.
- 29. Paluš, Milan, and Aneta Stefanovska. "Direction of coupling from phases of interacting oscillators: An information-theoretic approach." Physical Review E 67.5 (2003): 055201.
- Vejmelka, M. and Paluš, M., 2008. Inferring the directionality of coupling with conditional mutual information. Physical Review E, 77(2), p.026214.
- O'Reilly, R.C. and McClelland, J.L., 1994. Hippocampal conjunctive encoding, storage, and recall: avoiding a tradeoff. Hippocampus, 4(6), pp.661-682.
- Zhang, T., Zeng, Y., Zhao, D., Wang, L., Zhao, Y. and Xu, B., 2016, October. Hmsnn: hippocampus inspired memory spiking neural network. In 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 002301-002306). IEEE.
- Amaral, D.G., Scharfman, H.E. and Lavenex, P., 2007. The dentate gyrus: fundamental neuroanatomical organization (dentate gyrus for dummies). Progress in brain research, 163, pp.3-790.

- Swaminathan, A., Wichert, I., Schmitz, D. and Maier, N., 2018. Involvement of mossy cells in sharp wave-ripple activity in vitro. Cell reports, 23(9), pp.2541-2549.
- Henze, Darrell A., and György Buzsáki. "Hilar mossy cells: functional identification and activity in vivo." Progress in brain research 163 (2007): 199-810.
- Scharfman, H.E., 1995. Electrophysiological evidence that dentate hilar mossy cells are excitatory and innervate both granule cells and interneurons. Journal of neurophysiology, 74(1), pp.179-194.
- Diamond, S., Balvin, R.S., Diamond, F.R., 1963. Inhibition and choice: A neurobehavioral approach to problems of plasticity in behavior. Harper & Row Publishers.
- Fu, J., Wang, H., Gao, J., Yu, M., Wang, R., Yang, Z. and Zhang, T., 2017. Rapamycin effectively impedes melamineinduced impairments of cognition and synaptic plasticity in Wistar rats. Molecular neurobiology, 54(2), pp.819-832.
- Cooper L N, Bear M F. The BCM theory of synapse modification at 30: interaction of theory with experiment[J]. Nature Reviews Neuroscience, 2012, 13(11):798-810.
- Canolty, R.T. and Knight, R.T., 2010. The functional role of cross-frequency coupling. Trends in cognitive sciences, 14(11), pp.506-515.
- Krotov, D. and Hopfield, J.J., 2019. Unsupervised learning by competing hidden units. Proceedings of the National Academy of Sciences, 116(16), pp.7723-7731.
- 42. Cès, A., Burg, T., Herbeaux, K., Héraud, C., Bott, J.B., Mensah-Nyagan, A.G. and Mathis, C., 2018. Age-related vulnerability of pattern separation in C57BL/6J mice. Neurobiology of aging, 62, pp.120-129.
- Carasatorre, M., Ochoa-Alvarez, A., Velazquez-Campos, G., Lozano-Flores, C., Diaz-Cintra, S.Y. and Ramirez-Amaya, V., 2015. Hippocampal synaptic expansion induced by spatial experience in rats correlates with improved information processing in the hippocampus. PloS one, 10(8), p.e0132676.
- 44. Myers, C.E. and Scharfman, H.E., 2009. A role for hilar cells in pattern separation in the dentate gyrus: a computational approach. Hippocampus, 19(4), pp.321-337.
- Martinetz, T. and Schulten, K., 1991. A" neural-gas" network learns topologies. Artificial neural networks, pp.397-402.
- 46. Pontes-Filho, S., Lind, P., Yazidi, A., Zhang, J., Hammer, H., Mello, G.B., Sandvig, I., Tufte, G. and Nichele, S., 2020. A neuro-inspired general framework for the evolution of stochastic dynamical systems: Cellular automata, random Boolean networks and echo state networks towards criticality. Cognitive Neurodynamics, 14(5), pp.657-674.
- Chen, X. and Yang, T., 2021. A neural network model of basal ganglia's decision-making circuitry. Cognitive Neurodynamics, 15(1), pp.17-26.

- Hung, S.L. and Adeli, H. (1994b), "Object-Oriented Back Propagation and Its Application to Structural Design", Neurocomputing, Vol. 6, No. 1, pp. 45-55.
- 49. Adeli, H. and Hung, S.L. (1994), "An Adaptive Conjugate Gradient Learning Algorithm for Effective Training of Multilayer Neural Networks", *Applied Mathematics and Computation*, 62(1), pp. 81-102.
- 50. Myers, C.E. and Scharfman, H.E., 2011. Pattern separation in the dentate gyrus: a role for the CA3 backprojection. Hippocampus, 21(11), pp.1190-1215.
- Rolls, E.T., 1989. Functions of neuronal networks in the hippocampus and neocortex in memory. In Neural models of plasticity (pp. 240-265). Academic Press.
- 52. Rolls, E.T., 1996. A theory of hippocampal function in memory. Hippocampus, 6(6), pp.601-620.
- Tripp, B. and Eliasmith, C., 2016. Function approximation in inhibitory networks. Neural Networks, 77, pp.95-106.
- Person, A.L. and Perkel, D.J., 2005. Unitary IPSPs drive precise thalamic spiking in a circuit required for learning. Neuron, 46(1), pp.129-140.
- Mackwood, O., Naumann, L.B. and Sprekeler, H., 2021. Learning excitatory-inhibitory neuronal assemblies in recurrent networks. Elife, 10, p.e59715.
- Zhou, S. and Yu, Y., 2018. Synaptic EI balance underlies efficient neural coding. Frontiers in neuroscience, 12, p.46.
- Rubin, R., Abbott, L.F. and Sompolinsky, H., 2017. Balanced excitation and inhibition are required for highcapacity, noise-robust neuronal selectivity. Proceedings of the National Academy of Sciences, 114(44), pp.E9366-E9375.
- Hainmueller, T. and Bartos, M., 2020. Dentate gyrus circuits for encoding, retrieval and discrimination of episodic memories. Nature Reviews Neuroscience, 21(3), pp.153-168.
- Rafiei, M.H. and Adeli, H., 2017. A New Neural Dynamic Classification Algorithm, IEEE Transactions on Neural Networks and Learning Systems, 28(12), pp.3074-3083
- Pereira, D.R., Piteri, M.A., Souza, A.N., Papa, J., and Adeli, H., 2020. FEMa: A Finite Element Machine for Fast Learning. Neural Computing and Applications, 32:10, 6393-6404.
- Ahmadlou, M. and Adeli, H., 2010. Enhanced Probabilistic Neural Network with Local Decision Circles: A Robust Classifier, Integrated Computer-Aided Engineering, 17(3), pp. 197-210.
- Alam, K.M., R., Siddique, N., Adeli, H., 2020. A Dynamic Ensemble Learning Algorithm for Neural Networks, Neural Computing with Applications, 32(10), pp.8675-8690.