

Article

Not peer-reviewed version

Model of Non-Hebbian Engram Formation for Visual Recognition

[O.V. Levashov](#) and [V.F. Safiulina](#) *

Posted Date: 23 April 2026

doi: 10.20944/preprints202604.1673.v1

Keywords: top-down recognition model; neural network modeling; magnocellular system; matching; Hebbian principle; veto neurons; ripples; non-Hebbian learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Model of Non-Hebbian Engram Formation for Visual Recognition

O.V. Levashov¹ and V.F. Safiulina^{2,*}

¹ Brain & Body Development Centre, Dalat, Vietnam

² Institute of Physiology, Federal Research Center, Komi Scientific Center, Ural Branch of the Russian Academy of Sciences (IF FRC, Komi Scientific Center, Ural Branch of the Russian Academy of Sciences)

* Correspondence: Victoria.Safiulina@gmail.com

Abstract

A neural model for the formation of visual engrams is proposed, operating according to a non-Hebbian principle — specifically, through the enhancement of inhibitory synapses, up to and including the formation of veto synapses. The model relies on two hypothetical mechanisms: (1) rapid, repetitive reactivation ("ripple-reverberation") and (2) high-frequency synchronization enabling the activation of inhibitory synapses, which consequently become veto synapses. Through such learning, "neural locks" for familiar patterns are formed in memory. This model constitutes a component of a more general top-down model of visual recognition described previously (Levashov & Safiulina, 2025). The problem of processing activity patterns in living neural networks is discussed, as these patterns are not holistic but rather manifest as a mosaic of activated and non-activated neurons.

Keywords: top-down recognition model; neural network modeling; magnocellular system; matching; Hebbian principle; veto neurons; ripples; non-Hebbian learning

Introduction

At first glance, the problem of visual image processing in living neural networks may seem uncomplicated and tractable, particularly given the successful implementation of analogous tasks using artificial neural networks within the deep learning paradigm. However, a fundamental distinction exists between formal neural networks and their biological counterparts. Activation patterns in biological neural networks are not holistic in the literal sense but rather constitute a mosaic of activated and quiescent neurons. While such patterns can be processed by a set of neurons with receptive fields — effectively implementing a mathematical convolution operation over fragments of the input pattern, with subsequent transmission of convolution outputs to higher cortical areas — this approach inevitably leads to a loss of topological connectivity among the extracted features. More critically, the mechanisms by which one might manipulate such patterns as unified entities within a single neural layer (e.g., translation, scaling, or rotation of the pattern as a whole) remain obscure. Furthermore, due to recurrent excitatory connectivity, any pattern transiently "held" for processing at a given neural layer begins to rapidly dissipate, losing its initial spatial configuration (Agnes & Vogels, 2024; Levashov & Safiulina, 2025).

Compounding this difficulty is the current absence of a formal mathematical framework capable of describing the operation of such an extraordinarily complex structure as the brain's neural networks. From a technical standpoint, the challenge arises from its representation as a three-dimensional graph comprising an immense number of chemically and electrically coupled elements, each possessing unique activity dynamics determined by its threshold, the weights of thousands of synapses, and the influence of numerous other neurons within the graph.

Consequently, the primary approach to investigating the operational principles of such a complex system as the visual brain involves formulating hypotheses and conceptual models

grounded in physiological data, subsequently developing algorithms (prompts) based on these models, and conducting computer simulations following these prompts.

In our previous work, we developed neural models and performed computer simulations of two distinct visual processing mechanisms. The distinctive features of our models include:

1. Incorporation of inhibitory neurons, including "veto" interneurons, into the network architecture, enabling the construction of neural circuits capable of executing sophisticated processing functions (Levashov & Safiulina, 2025).
2. Entrainment of the entire neural network via endogenous rhythms analogous to EEG rhythms, serving as "clock frequency" akin to those in electronic systems (Levashov & Safiulina, 2026).

In the present article, we propose a neural model for the formation of visual engrams according to a non-Hebbian principle — specifically, via potentiation not of excitatory, but of inhibitory synapses. To implement such learning, we introduce two key hypothetical mechanisms: (1) rapid, iterative repetition of the input pattern ("ripple reverberation"), and (2) rhythmic synchronization to engage inhibitory synapses, which consequently become "veto synapses" (gates). This model constitutes a component of a broader top-down model of visual recognition previously described by us (Levashov & Safiulina, 2025).

Modeling the Process of Visual Recognition in Living Systems

The challenge of modeling the ability of the visual brain (in humans and higher animals) to perceive, analyze, and comprehend the surrounding world remains one of the most difficult and unresolved problems to this day. For instance, it has not yet been reliably established whether any characteristic features are extracted prior to final recognition, and if so, which ones. Another major obstacle is determining exactly how the image being processed in the visual system is matched to templates (engrams) stored in memory.

Overall, three main approaches can be identified in the history of modeling visual recognition in living systems.

1. Mapping (or Matching). This is the idea of directly comparing a cortical copy of the retinal image with an engram in the cerebral cortex. The pioneer of this approach was Donald Hebb (Hebb, 1949). Hebb proposed that during learning, when a visual pattern is presented at the input, there will always be a group of neurons that fire together (Hebb, 1949). According to Hebb, if these neurons are identified and their connections strengthened, such neuronal ensembles would serve as templates ("engrams") for comparison with input representations.

However, it eventually became clear that the idea of transmitting the retinal image to the cortex for comparison with templates is unrealistic. The primary reason is that it is impossible to separate "figure" from "ground" at the retinal level, since the recognition process has not yet begun; consequently, obtaining a "pure" neural representation of the target object is impossible.

This led specialists to a new idea: to extract several characteristic fragments from the initial image, i.e., to extract "local features."

2. Feature-Based Recognition. The visual system does indeed possess the ability to extract local shape features. For example, neurons in visual cortical areas V1, V2, and V4 extract edges, corners, line endpoints, local color, and texture fragments. These features are encoded by the activity of various populations of cortical neurons (Hubel, 1988). Following the discovery of such "local form detectors" by neurophysiologists, several works on feature extraction models emerged. This approach is most explicitly articulated in David Marr's model (Marr, 1982). Marr proposed the so-called "computational theory" of vision. According to this theory, there are no templates in memory in the Hebbian sense; rather, there are sets of structured descriptions in the form of features and their spatial relationships.

Another widely cited model of visual recognition is Biederman's recognition-by-components model (Biederman, 1987). Biederman argued that object recognition in humans occurs through the mental decomposition of objects into elementary geometric primitives—geons (cylinders,

parallelepipeds, cones, spheres, wedges). He believed that combinations of 2–3 geons are sufficient for the unique description of most familiar objects. Recognition proceeds by matching the structural model (set of geons) extracted from the image with templates stored in memory. In this respect, the model is similar to Marr's. However, neither of the models described above has demonstrated effectiveness as a viable algorithm for visual recognition.

The difficulty with the feature-based recognition approach lies, in our view, precisely in the fact that the extracted features carry no information about their localization and thus are merely "lists." This is why, in the two aforementioned models, Marr and Biederman had to postulate a hypothetical "binding" mechanism. Furthermore, with this approach, the visual system inevitably encounters a combinatorial explosion: from a limited set of simple features (lines, angles), an infinite number of objects can be composed. Therefore, lists of features do not help solve the problem of object recognition in complex visual scenes. Moreover, physiological evidence indicates that different features of the same object (form, color, motion) are processed in different areas of the visual cortex. No single "screen" where the complete picture is assembled has been found in the brain.

3. Predictive ("Top-Down") Recognition Models. Another approach to understanding recognition mechanisms involves predictive models, in which rapid but coarse (based on overall shape similarity) recognition occurs first—i.e., a rapid hypothesis selection—followed by its verification.

The most physiologically grounded model of this type is described in the work of Bar, Kveraga, and colleagues (Bar, 2003; Kveraga et al., 2007). Although purely descriptive, this model was formulated based on fMRI and MEG data obtained by this research group from subjects performing a visual recognition task. This allows the use of this "top-down" scheme as a basis for further modeling of specific processing mechanisms for visual recognition.

According to the Bar et al. model (Bar, 2003; Kveraga et al., 2007), low-frequency information in the object image (silhouettes, blobs) is transmitted to the prefrontal cortex via the magnocellular pathway, arriving there earlier than the more detailed image containing color and texture components, which is transmitted via the parvocellular pathway. Based on the integration of the "coarse" and more detailed descriptions, the prefrontal cortex forms an "initial guess." This guess is then sent "down" to area V4, suppressing neurons that signal irrelevant features. Simultaneously, the prefrontal cortex issues a command to the superior colliculi, which control saccadic eye movements, to shift gaze to another location in the scene where, according to the hypothesis, a predicted shape feature should be located. For example, "if this is a cat, then there should be an ear there." If, after the saccade, the "ear" is indeed recognized, the initial hypothesis is confirmed.

Evidence supporting this recognition paradigm was also obtained by Malcolm and Henderson (Malcolm & Henderson, 2009). Subjects had to find a specific object in a photograph of a complex scene (an interior). The target was specified either verbally (the word "clock" was spoken) or visually (a photograph of the exact clock to be found). The specificity of the cue did not affect the latency of search initiation but significantly influenced the scanning and verification stages. It turned out that with a specific visual template, participants scanned the scene more quickly and made faster decisions regarding target detection. Furthermore, when scanning the scene with a visual cue, participants fixated on fewer regions of interest (ROIs), and the average duration of their fixations was shorter. Importantly, even with a specific template, scene scanning required several saccades. Levashov's work (Levashov, 2018) also recorded subjects' saccades after brief presentations of complex, colored, multi-figure scenes. It was found that the very first saccade was directed not randomly, but to a meaningful location in the scene, even though subjects could report nothing about the scene's content. This may indicate the automatic generation of a hypothesis based on incomplete data. Typically, at least three to four saccades were required to "comprehend" the scene's gist.

Within the "top-down" framework, the first part of our recognition model has also been developed (Levashov & Safiulina, 2025), briefly described below.

Magnocellular Model of Visual Recognition

Our proposed model is part of a more general concept of visual recognition developed by the author and colleagues (Pozin et al., 1978; Levashov, 2018, 2022). According to this concept, the visual brain comprises two main blocks for recognizing objects in the environment: one associated more with the left hemisphere (LH) cortex, and the other with the right hemisphere (RH) cortex.

The recognition block in the right RH differs little from the visual brain of higher animals, whose primary function during evolution was to ensure survival. Consequently, visual recognition there is directly linked to the amygdala and hippocampus, which treat all objects in the environment as dangerous, attractive, or neutral. The visual brain of an infant, who has not yet formed concepts about objects, operates similarly. However, during the maturation of the frontal cortex, a second block, more associated with the left LH—the categorical recognition block—also forms in the visual brain. Its connection to the left LH is obvious, as it is in the left LH that centers for speech comprehension and the understanding of words as labels for visual objects and their uses develop.

It is pointless to determine which of the two blocks is primary, as the human brain solves a vast number of tasks, and the involvement of the "left" and "right" blocks may differ for each task. For example, during reading, the right block is responsible for the holistic perception of paragraphs, lines, sentences, and words, while the left block is responsible for recognizing individual letters. At the same time, the temporal cortex of the right RH specializes in recognizing faces of close and familiar people (Kok, 2022).

Over time, it became clear that the visual brain has another dichotomy: magno- and parvocellular conducting neural channels (the dorsal and ventral streams), which begin in the retina and pass through the LGN to higher cortical areas of the hemispheres (Zeki, 1993). The magnocellular system (M-system), being faster in terms of signal conduction latency, can be considered a structure for preliminary recognition, while the P-system, carrying information about color and fine details, can be considered a structure for the final stage of recognition (Zeki, 1993).

Our proposed model (conditionally designated as the M-model) corresponds to the preliminary recognition block in the left LH cortex.

The stages of the M-model's operation are as follows:

1. Defocusing (blurring) of the portion of the visual scene selected by central vision.
2. Binarization and transformation of the image into a two-level "silhouette" that preserves the topology of the recognized object's shape.
3. Mapping (matching by superimposition) with contextually relevant templates in memory.
4. Selection of a single template based on the "winner-take-all" principle.
5. The selected shape, as a visual hypothesis, is transmitted "down" to lower levels of the visual system and visualized.
6. Thereby, the brain (frontal cortex) decides that the object is identified.
7. Hypothesis verification is performed using subsequent saccades to presumed "areas of interest."

This procedure allows for a rapid decision after the very first fixation, enabling danger avoidance in critical situations, and, in the absence of danger, allows for hypothesis verification and completion of recognition.

This model is based on the following physiological data:

1. The M-system is known to dominate in the left LH (Okubo & Nicholls, 2005). This study found a left-hemisphere advantage in the temporal resolution of visual stimuli, indicating the key role of the M-system in interhemispheric asymmetry during visual information processing.
2. Receptive fields of M-neurons in the retina and LGN have a concentric structure and large dimensions; therefore, from a technical standpoint, they perform an image blurring (low-pass filtering) operation (Zeki, 1993; Bar et al., 2003).
3. According to Bullier (Bullier, 2001), the pattern of activity from the retina is transmitted via the LGN directly to the prefrontal cortex, bypassing layers of "simple form detectors." Signals from the magnocellular layers of the LGN reach the cortex 10–20 ms faster than parvocellular signals, and the first responses in area MT (V5) appear approximately 35–45 ms after image presentation to the retina. Importantly, this bypasses neurons with receptive fields tuned to lines, angles, and

other forms, which ensures the preservation of the processed pattern's topology for mapping (see discussion of this issue in the Introduction).

4. Adusei et al. (Adusei et al., 2024) found that feedback from areas V4 and MT to the macaque lateral geniculate nucleus is organized as parallel streams, specifically activating certain thalamic layers. This suggests the ability of higher visual centers to directly modulate visual information transmission, i.e., it supports the "top-down" paradigm of visual recognition.
5. According to M. Bar and colleagues, who studied the recognition process using combined fMRI and MEG measurements, the first area activated after image presentation to the retina is the orbitofrontal cortex of the left LH (latency 130 ms), followed by the temporal cortex of the right LH (latency 180 ms).

Described below is the second part of our M-model – a model for engram formation to perform matching at the final stage of recognition.

Non-Hebbian Learning Principle in Engram Formation

One important task for the visual brain is the visual search for an object in a scene based on a template. Humans often solve this task in their activities (e.g., an assembler, a shopper in a store, a child building a house from construction set pieces). This task was investigated by Malcolm and Henderson (Malcolm & Henderson, 2009). Subjects were shown photographs of real scenes (e.g., a kitchen interior), and their eye movements were recorded during target search. The search target (e.g., "clock") was specified either verbally or visually – by a photograph of the exact clock to be found. The presence of a visual template allowed participants to scan the scene significantly faster and make faster decisions when the target was already located but not yet identified.

The classic solution for forming an object engram (template) in memory is Hebb's model of synaptic plasticity (Hebb, 1949). According to Hebb's rule, "neurons that fire together, wire together." It is believed that this is how templates (engrams) are formed in visual memory. The Hebbian doctrine has been criticized repeatedly. For instance, Rochester et al. (Rochester et al., 1956) were among the first to empirically demonstrate that without weight decay, a network formed according to Hebb's rules saturates quickly. Grossberg (Grossberg, 1976) also strongly criticized the Hebbian model for failing to solve the "stability-plasticity dilemma."

The main criticism of Hebb's rule is that it does not provide a decision criterion. Hebb proposes a mechanism for "recording" information at the synaptic level but does not explain how, from a multitude of neurons, one subsequently selects the ensemble that best matches the target.

In this article, we propose a neural model of learning and formation of "templates" in the visual system's memory, based on a principle different from the classic Hebbian principle. In our model, during learning, the weights of connections increase not between co-active excitatory neurons (Hebbian principle), but between active and inactive neurons. Moreover, these connections are inhibitory and, by the end of learning, become "veto connections" (prohibitory). A rapid "reverberation" stage is also provided, during which relevant inhibitory connections become "veto" connections. As a result, a kind of neural "locks" are formed, which allow "self" patterns to pass through and block "foreign" patterns.

Let us examine our proposed principle of template (engram) formation in more detail.

According to the classic Hebbian rule, "neurons that fire together, wire together," and consequently, the weights of synaptic connections between them increase (Hebb, 1949). In contrast to the "Hebbian" rule, we use "non-Hebbian rules" (Figure 1):

1. If one neuron is activated and a neighboring neuron is not activated, the activated neuron forms (activates) an inhibitory synapse on the neighboring neuron. Upon repeated activation, the synaptic weight increases by a small amount (δ).
2. The process of memorizing the input pattern (learning) occurs during "reverberation" – rapid repetition of the current pattern at the input. We hypothesize that the frequency of such

repetitions is set by an external rhythm, such as a "ripple" (150-200 Hz), observed during memory consolidation in the hippocampus and prefrontal cortex.

- As a result of learning, the weights of the inhibitory synapses become so large that they can be considered "veto synapses" (prohibition), blocking signal transmission along the corresponding pathway.
- This learning procedure yields a kind of "imprint" of the current pattern in the form of an "individual lock." Subsequently, such a "lock" is "opened" only upon the arrival of the "self" pattern or a part thereof at the input.

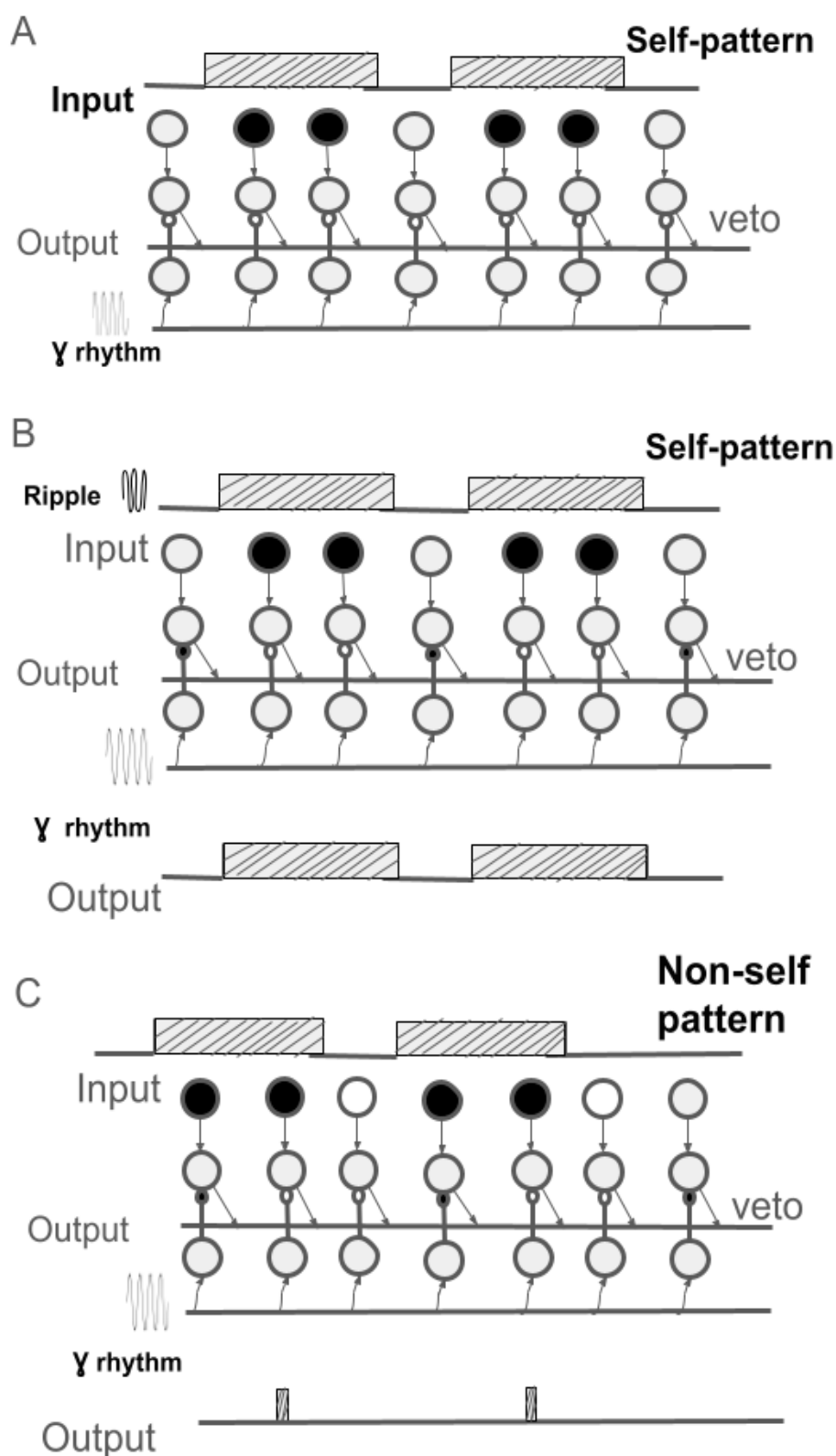


Figure 1. A. Beginning of learning (pattern arrives at network input, top). Neurons activated by the pattern are shown in black. Simultaneously with pattern presentation, a high-frequency rhythm (gamma rhythm in this example) is applied to the "bus" at the bottom. This initiates the interaction of signals within this network. According to the proposed algorithm, activated neurons from the pattern activate neurons in the layer below, while simultaneously, neurons in the lowest layer are activated from the bus. Inhibitory synapses from layer 3 neurons onto activated layer 2 neurons remain unchanged. Meanwhile, on the inactivated neurons of layer 2, under the influence of "ripple-reverberation," inhibitory synapses with large weight, i.e., "veto-synapses," gradually form, endowing their "veto-neurons" with the property of blocking signal transmission along that pathway. B. The state of the neural network after learning with "ripple-reverberation." Black indicates "veto" type synapses formed during learning (20 pattern repetitions in computer modeling). If the "self" pattern reappears at the input, it "passes through" the formed "neural lock." C. The same neural network after learning. For simplicity, the same pattern is presented at the input, but shifted. Such a "foreign" pattern does not pass through the formed "veto pathways" and is effectively sorted out (only a negligible residue passes through).

To test the functionality of this model, a Python program was used. A detailed description of the methodology and results of computer modeling of matching for objects in the form of conditional test "letters" is provided in the Appendix.

Figure 2 shows examples of matching simulation for test patterns. This figure shows examples of patterns that passed through the "locks" of template "letters," as well as blocked patterns.

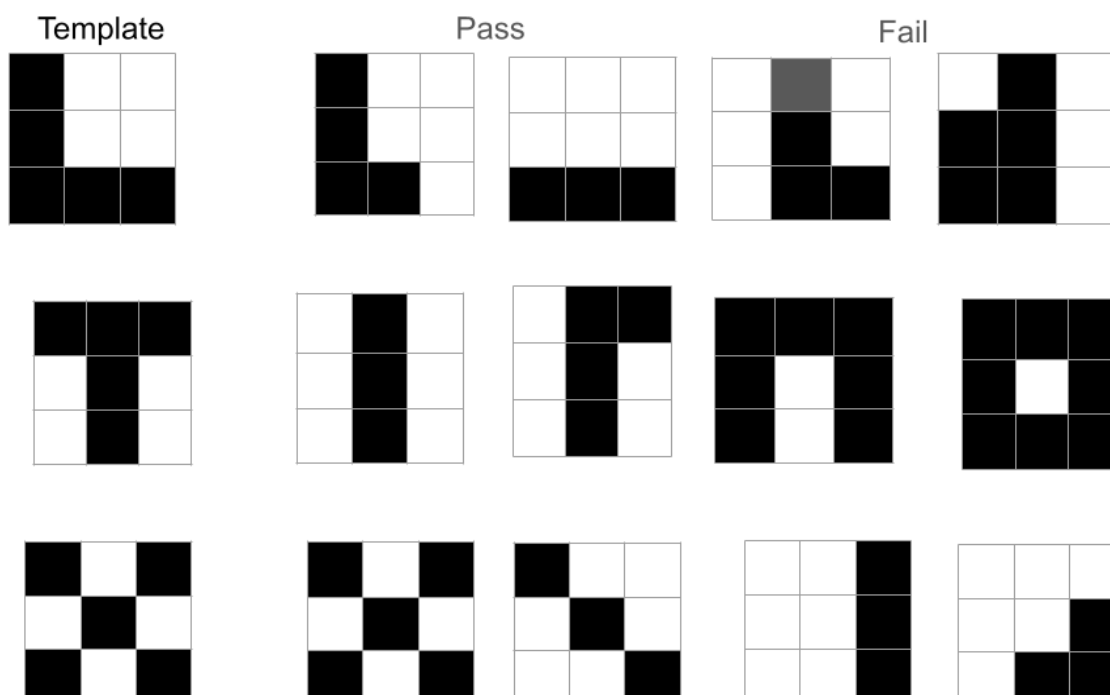


Figure 2. Example of matching simulation after template formation for "letters." The left column shows the templates. The second column shows the configurations that pass through the formed "template-locks." The third column shows blocked patterns. It can be seen that some "foreign" patterns pass through the "locks" formed during learning; however, they are nothing other than parts of the original template. In this case, the program considers them similar to the "self pattern." We note that in such a model, large parts of the original templates can be considered characteristic "features" for recognition. In this case, such a "feature" can generate a hypothesis about the presence of a familiar pattern in a given location of the visual scene. Thus, even partial matching between the input pattern and the template allows initiating the hypothesis verification procedure described above in the section on the top-down model of visual recognition.

Computer modeling of matching on this set of templates and patterns showed that no "foreign" pattern was mistakenly accepted as "self." At the same time, proper fragments were recognized as "familiar."

Conclusions from Computer Modeling Results

1. The modeling demonstrated that the procedure for template (engram) formation proposed in this article solves this problem under the condition of two additional mechanisms: rapid repeated pattern presentation (ripple-reverberation) and the presence of a high-frequency rhythm (e.g., gamma rhythm) for activating key elements of the neural network – the inhibitory synapses that become "veto-synapses" (prohibitory) at the end of learning.
2. A feature of this procedure is that the formed local neural network "passes" both the entire pattern and its large fragments (parts of the whole). Such fragments can, in essence, be considered local features. Thus, our proposed M-model of recognition acquires a new form – mapping combined with local feature extraction.
3. The ability to generate plausible hypotheses based on a fragment of the whole pattern implies good noise immunity of such an algorithm under conditions of visual noise (interference), as well as when analyzing real complex scenes where object occlusion (partial overlap of the contours of the target object and distractors) is common.

General Discussion

Our proposed model of template formation is grounded in real physiological data.

Veto Neurons. One of the earliest mentions of "veto" neurons was the article by Trevelyan and colleagues (Trevelyan et al., 2006). The study showed that the spread of epileptiform activity in the mouse neocortex is regulated by "vetoing inhibition," restraining the firing of pyramidal neurons. The first direct demonstration of "inhibitory veto" in the neocortex (including visual cortex V1) *in vitro* showed that local "veto" interneurons block spike propagation. The presence of different types of inhibitory neurons in the cortex, including "veto" neurons, has been shown in many articles, including that of Pfeffer et al. (Pfeffer et al., 2013). This study identified three types of interaction among three major interneuron populations in the mouse visual cortex: 1. Parvalbumin-expressing interneurons, which strongly inhibit each other but weakly inhibit other populations. 2. Somatostatin-expressing interneurons, which do not inhibit each other but strongly inhibit all other neuron types. 3. Vasoactive intestinal peptide-expressing interneurons, which primarily inhibit somatostatin interneurons. Elstrott and Feller (Elstrott & Feller, 2009) detail the mechanisms of lateral inhibition, which serve as the cellular basis for suppressing responses to distracting stimuli – a functional analogue of "veto." Bergoin et al. (Bergoin et al., 2023) examined various types of interactions between excitatory and inhibitory neurons and concluded that inhibitory neurons act as "stabilizers" in plastic neural networks, ensuring long-term retention of information acquired during learning. The "veto" concept also applies at higher levels, where neurons selective for certain features can suppress responses to irrelevant stimuli.

Ripples (Sharp-Wave Ripples, SWRs). These are high-frequency oscillations (140-200 Hz) in the hippocampus, considered a neurophysiological correlate of memory consolidation. Traditionally studied in the hippocampus, they are believed to provide "replay" (reactivation) of neural ensembles encoding previous experience and facilitate the transfer of this information to the neocortex for long-term storage. However, it was unknown whether similar events are generated directly in associative cortical areas (neocortex) and how exactly the cortex and hippocampus interact during learning. Khodagholy et al. (Khodagholy et al., 2017) discovered that during sleep, ripples occur not only in the hippocampus but also in the association cortex, and their synchronization is enhanced after learning. Using flexible organic transistors (NeuroGrid) for recording, the authors simultaneously recorded activity from the cortical surface (electrocorticogram, ECoG) and deep structures (hippocampus) in freely behaving rats. They discovered high-frequency oscillations similar to

hippocampal ripples in associative cortical areas (specifically, the medial prefrontal, perirhinal, and entorhinal cortex). The frequency of these cortical events was lower (100-140 Hz) than in the hippocampus, and they were generated locally, not merely as a result of signal transmission from the hippocampus. During active behavior, hippocampal and cortical ripples were largely independent and not synchronized. However, after performing an attention-demanding task, the interaction pattern changed. During subsequent rest or sleep periods, a significant increase in the number of co-occurring (coherent) ripple episodes was observed, where a hippocampal ripple occurred simultaneously or with a short delay relative to a ripple in the association cortex. Temporal delay analysis suggested that in these co-occurring events, the hippocampal ripple typically slightly led the cortical ripple, supporting the hypothesis that the hippocampus "leads" the cortex, transferring information for consolidation. More recent work, e.g., by Mishra et al. (Mishra et al., 2024), examined patients with implanted electrodes, showing that hippocampal SWRs co-occur with cortical ripples (especially in the anterior temporal lobe), concluding that this interaction coordinates the retrieval of semantic information from memory. Robinson et al. (Robinson et al., 2025) identified a subgroup of "large" (in amplitude and extent) SWRs, arguing that these, rather than ordinary ripples, are critically important for reactivating neural ensembles in the prefrontal cortex and for successful skill consolidation during sleep. Iwata et al. (Iwata et al., 2024) found that the incidence of SWRs in humans increases during moments of spontaneous thoughts and recollections ("self-generated thoughts") unrelated to the current external task, confirming the role of SWRs in supporting the internal mental stream. Yang et al. (Yang et al., 2024) investigated how the brain "decides" what needs to be remembered, noting SWRs occurring immediately after an event (during wakefulness), which they suggest tag neural sequences for subsequent replay and consolidation during sleep. Finally, Janssen et al. (Janssen et al., 2025) used fiber photometry and electrophysiology to show that shortly after SWR occurrence in the hippocampus, there is a dopamine surge in the nucleus accumbens, indicating that the brain "rewards" itself for replaying important experiences, crucial for reward-based learning.

High-Frequency Synchronization. Unlike artificial neural networks, the pattern of excitation and inhibition in a real neuronal layer is not an integrated pattern but rather a "mosaic" of activated and inactivated pixels (speaking technically). Therefore, it was unclear how such a mosaic could be manipulated – e.g., shifted or rotated – although such manipulations are theoretically characteristic of visual recognition. Previously, we demonstrated that the presence of high-frequency synchronization (a kind of clock frequency) allows simulating the integrated processing of such a mosaic pattern (Levashov & Safiulina, 2025, 2026).

Application

The proposed computational neural network model implements a mechanism for discriminating input patterns based on their correspondence to previously formed memory engrams. The network architecture comprises three layers: an afferent input layer, a modulated template layer, and a system of inhibitory "veto interneurons" that control the output signal. Input activity is modeled as spatially distributed impulses, which are compared during processing with internal standards ("self-patterns"), represented as fixed synaptic conductance zones. The model's operation is strictly synchronized in time by two types of oscillatory activity: the gamma rhythm and the high-frequency ripple complex. The gamma rhythm activates inhibitory neurons, while the ripple reverberation simulates a state characteristic of the consolidation phases or data retrieval from memory. The system's reliability is tested by presenting "foreign patterns." Detailed implementation of the network for modeling on a 3x3 matrix. The study involved computer modeling the process of storing a set of different visual stimuli into memory. Letter patterns (e.g., "T," "L," "P," "X," "C," "O") inscribed in a 3x3 square matrix were used as search objects. (Figure 3)

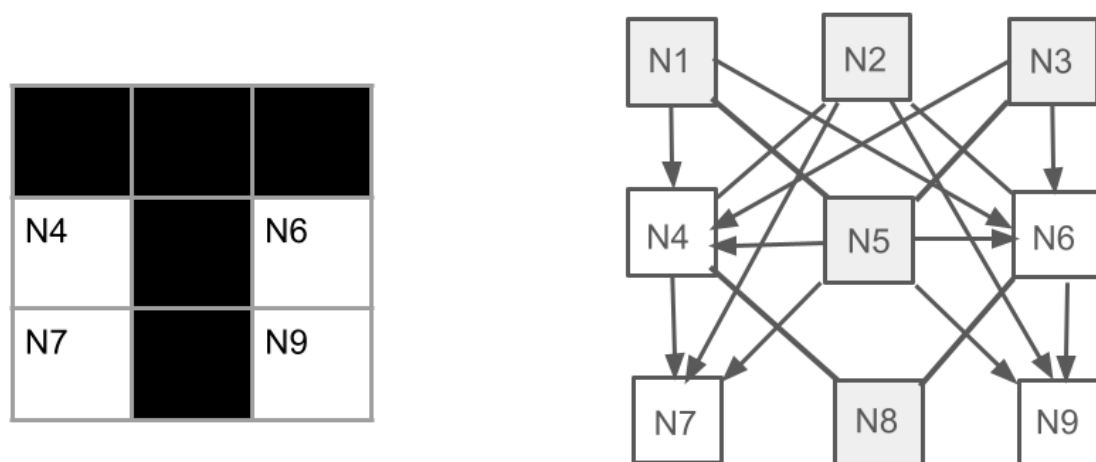


Figure 3. Architecture and connection topology of the 3x3 neural network during training on the reference pattern (the letter "T"). The matrix representation of the input image is shown on the left. Active neurons ($N1$, $N2$, $N3$, $N5$, $N8$) that form the structure of the recognizable symbol are highlighted in black. White cells denote inactive (background) elements of the matrix ($N4$, $N6$, $N7$, $N9$). The neural network graph after completing the training cycle is shown on the right. Gray nodes correspond to active points of the pattern. Synaptic connections are thick lines (or connections between black nodes): Excitatory connections with positive weight, established between adjacent active neurons to ensure the integrity of the image. Thin lines with arrows: inhibitory connections ("inhibitory synapses") with negative weight, directed from the active neurons of the pattern to the background nodes. This configuration of weights minimizes the probability of false activation of the network when a noisy input signal is applied and ensures the stability of the recognition of the reference vector.

Each of the nine neurons of the matrix corresponds to one pixel of the binary image and can be in one of two states: active (value of 10 conventional units) or inactive (0). The activation threshold of a neuron is taken to be 5, which ensures reliable separation of the signal from the background. Before training, a basic 9x9 weight matrix W is formed (connections between all pairs of different neurons). All weights are initialized to the value of -1.0 , which corresponds to weak background inhibition. Diagonal elements (connections of a neuron with itself) are absent. Engram formation algorithm ("non-Hebbian" learning). The process of memorizing the reference pattern occurs in a single presentation cycle with subsequent weight fixation. The following steps are performed for each reference: 1. Formalization of the stimulus. Each letter is represented as a state vector of nine neurons. Active elements (black cells in the image) are assigned a high signal level (10 conventional units), inactive ones - zero. The vector determines the set of active neurons (activity above the threshold) and the set of inactive neurons (activity below the threshold) 2. Topology analysis. The system identifies pairs of "active neighbors" (connections within the letter contour), which retain their initial synaptic weights (-1.0) during the training process. Neighborhood is determined based on the four-connectivity principle (up, down, left, right) in a 3x3 matrix. 3. Generation of "veto connections". The key stage of training is the activation of the "non-Hebbian" mechanism. From each neuron included in the structure of the letter, directional inhibitory connections (veto synapses) are formed to all neurons that should remain passive in this pattern. The weight of each such connection changes from -1.0 to -50.0 . This value is chosen so that even a single impulse from a forbidden neuron is guaranteed to block the signal from passing through the network. 4. Creation of a "neural lock". As a result, a unique configuration of inhibitory connections is formed for each letter in memory, blocking any neuronal activity that extends beyond the reference contour. Connections between active neurons remain unchanged (-1.0), preserving the integrity of the image. This algorithm differs fundamentally from Hebb's: it is not the connections between simultaneously active elements that are strengthened, but the connections from the active element to the inactive one. This creates the "lock" effect, which can only be opened by the pattern (or part of it) that does not contain "forbidden" elements. Testing procedure (matching) Matching modeling consists of passing input signals through the formed

"locks." The testing algorithm is configured for high selectivity with respect to the object's topology. Binary activity is calculated for the test vector. A pattern is considered "native" (passes through the lock) if the condition is met that all active neurons in the test vector belong to the set of active neurons in the reference vector. In other words, the test vector must be a strict subset of the reference vector. If at least one active neuron that was absent from the standard is detected in the test, such a neuron is affected by the "veto-synapse", and the resulting network output is zero — the pattern is blocked. The main feature of the method is that the recognition system classifies the input pattern as valid not only when all elements completely match, but also when the presented configuration is a strict subset of the standard. Thus, any part of the original structure that has preserved the internal geometry of the "lock" is recognized by the system as "similar" to the original, which ensures the stability of recognition with incomplete data. Stimulus material. Six standard patterns were selected for modeling, corresponding to the letters L, T, X, C, O, P, inscribed in a 3x3 matrix. Each standard represents a unique combination of active neurons: - L: N1, N4, N7, N8, N9 (5 active, 4 inactive); - T: N1, N2, N3, N5, N8 (5 active, 4 inactive); - X: N1, N3, N5, N7, N9 (5 active, 4 inactive); - C: N1, N2, N3, N4, N7, N8, N9 (7 active, 2 inactive); - O: N1, N2, N3, N4, N6, N7, N8, N9 (8 active, 1 inactive); - P: N1, N2, N3, N4, N6, N7, N9 (7 active, 2 inactive). To demonstrate the selectivity of the network, three groups of test patterns were prepared: - Inversions of patterns (active and inactive cells change places); - Competing shapes (for example, the letter H in relation to T); - Partial fragments of patterns (vertical line from L, upper crossbar from T, diagonal from X, etc.). All test patterns are encoded similarly to the reference patterns (10 – active, 0 – inactive). Results Formation of a "neural lock" using the "T" reference pattern as an example Figure 3 shows the result of training the network on the reference pattern corresponding to the letter "T". The active neurons in this pattern are N₁, N₂, N₃, N₅, N₈ (upper horizontal and central vertical). The remaining four neurons (N₄, N₆, N₇, N₉) should remain inactive. During training, each of the five active neurons established veto synapses with a weight of -50.0 on all four inactive neurons. The total number of veto connections formed was 20. The connections between the active neurons themselves retained their original weight of -1.0, which ensures the possibility of joint activation of all elements of the letter as a whole. Thus, a unique configuration of inhibitions was created in memory for the "T" standard: any attempt to activate, for example, neuron N₄ (upper right corner) upon presentation of a test pattern will immediately trigger inhibition from all neurons active in the standard, and the signal will not reach the output. The response of the network trained on the letter "T" was compared with responses to various test patterns (Figure 3). The network reliably passes not only the complete standard but also any of its fragments, for example, only a horizontal line (N₁, N₂, N₃) or only a vertical line (N₂, N₅, N₈), which corresponds to the property of partial recognition. At the same time, any pattern containing at least one neuron absent from the standard is blocked. Thus, the letter "L" activates neurons N₄, N₇, N₉, which are prohibited and therefore does not pass. The reaction to the "T" inversion (N₄, N₆, N₇, N₉ are active) is particularly indicative: none of the active neurons of the inversion are included in the standard, so they are all under veto prohibition, and the pattern is blocked. Quantitative characteristics of veto connections For each standard, the number of formed veto synapses was counted. It is equal to the product of the number of active neurons of the standard and the number of inactive ones. For the L, T, X standards (five active, four inactive), the number of veto connections was 20. For the C and P standards (seven active, two inactive) - 14. For the O standard (eight active, one inactive) - 8. The number of veto connections determines the degree of "rigidity" of the lock: the more prohibitions, the more difficult it is for a random pattern to pass through the network. Patterns with a large number of inactive neurons (L, T, X) generate the maximum number of inhibitions and, therefore, exhibit the highest selectivity. Pattern O has only 8 veto connections, which, however, is sufficient to block any patterns that activate its single inhibited neuron (for example, the inversion of O, which activates only the center). Subset Principle Testing For each template, all possible fragments (subsets of active neurons) were tested (Figure 3). In all cases, the fragments successfully passed through the network. This proves that the proposed learning mechanism preserves information not only about the complete image, but also about all its constituent parts, which is an important property for visual

recognition under partial occlusion or incomplete field of view. Summary The results fully confirm the performance of the proposed non-Hebbian algorithm for engram formation. The 3x3 network, trained on six different letter patterns, demonstrated: 1) High selectivity: no outlier pattern (including inversions and competing forms) was mistaken for the template. 2) Robustness to incomplete data: any fragment of the template is recognized as "one's own." 3) Quantitative predictability: the number of veto connections directly depends on the number of inactive neurons in the template. Thus, the model successfully implements the "neural lock" principle, which can serve as the basis for top-down mechanisms of visual recognition, where rapid matching of the input silhouette with the standards stored in memory is required.

Conclusions

1. The difficulty of manipulating activity patterns in real neural networks is discussed, as they are not holistic patterns but merely a mosaic of activated and inactivated neurons. A solution is proposed in the form of local synchronization of network neuron activation, which keeps the processing sequential but makes it very fast.
2. A neural model of "non-Hebbian" engram formation in memory is proposed, based on the strengthening of inhibitory synapses up to a "veto" (prohibition) state. As a result, a kind of "neural locks" for "self" patterns are formed in memory.
3. The task of retaining the required engram in memory is solved by means of hypothetical "ripple-reverberation" – rapid repeated presentation of the to-be-remembered pattern.
4. Computer modeling demonstrated that the formed "lock"-type engrams "pass" (respond to) both the entire pattern and its large fragments (parts of the whole). Such fragments can be considered local features. Consequently, this recognition model acquires a new form – as a variant of the matching model incorporating local feature extraction.
5. The ability to generate plausible hypotheses based solely on a fragment of the target pattern implies good noise immunity of such a scheme under conditions of visual noise, as well as when analyzing real complex scenes with occlusion (where object contours partially overlap).
6. The plausibility of this model of engram formation for visual recognition could be tested in experiments involving rapid presentation of a fragment of a meaningful shape against random noise. According to our prediction, subjects would report perceiving the entire shape completely.

Author Contributions: O.L. – research concept, neural models, writing text, preparing illustrations; V.S. – modeling, editing, preparing illustrations.

Funding: The work had no financial support.

Code Availability: A software implementation of the neural network model (Python 3.x) with non-Hebbian learning, veto synapse formation, and a matching procedure for 3x3 patterns is available in the public repository: <https://github.com/victoriasafiulina-design/non-hebbian>.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Kok, E. P. Visual agnosia: syndromes of disorders of higher visual functions in unilateral lesions of the temporo-occipital and parietal-occipital regions of the brain. Moscow: URSS: LENAND Publishing House. 2022. 224 p.
- Levashov O. V. Attractors of visual attention and analysis of visual scenes. *Sensory Systems*. 2018. 32(3). 200–209.
- Levashov O.V. Artificial Vision. *Artificial Intelligence. Neural Models of Living Sensory Systems*. Moscow: URSS: LENAND Publishing House. 2022. 246 p.

- Levashov O. V., Safiulina V. F. Models of visual recognition and the problem of constancy of the shape of a visual pattern during its processing in the cortex. *Asymmetry*. 2025. 19(1): 6–17.
- Pozin N.V., Lyubinsky I.A., Levashov O.V. and others (edited by Pozin N.V.) *Elements of the theory of biological analyzers*. Moscow: Nauka.1978, 360 p.
- Adusei M, Callaway EM, Usrey WM, Briggs F. Parallel Streams of Direct Corticogeniculate Feedback from Mid-level Extrastriate Cortex in the Macaque Monkey. *eNeuro*. 2024.11(3):ENEURO.0364-23.2024.
- Agnes, E.J., Vogels, T.P. Co-dependent excitatory and inhibitory plasticity accounts for quick, stable and long-lasting memories in biological networks. *Nat Neurosci*. 2024. 27:964–974.
- Bar M. A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci*. 2003.15(4):600-9.
- Bergoin R, Torcini A, Deco G, Quoy M, Zamora-López G. Inhibitory neurons control the consolidation of neural assemblies via adaptation to selective stimuli. *Sci Rep*. 2023.13(1):6949.
- Biederman I. Recognition-by-components: a theory of human image understanding. *Psychol Rev*. 1987.94(2):115-147.
- Bullier J. Integrated model of visual processing. *Brain Res Brain Res Rev*. 2001 .36(2-3):96-107.
- Elstrott J, Feller MB. Vision and the establishment of direction-selectivity: a tale of two circuits. *Curr Opin Neurobiol*. 2009.19(3):293-7.
- Grossberg S. Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biol Cybern*. 1976. 23(3):121-34.
- Hebb, D. O. 1949. *The Organization of Behavior: A Neuropsychological Theory*. Wiley.
- Hubel, D.H., 1988. *Eye, brain, and vision*. New York: Scientific American Library.
- Iwata T, Yanagisawa T, Ikegaya Y, Smallwood J, Fukuma R, Oshino S, Tani N, Khoo HM, Kishima H. Hippocampal sharp-wave ripples correlate with periods of naturally occurring self-generated thoughts in humans. *Nat Commun*. 2024. 15(1):4078.
- Janssen MA, Chen HT, Tritsch NX, van der Meer MAA. Ventral Striatal Dopamine Increases following Hippocampal Sharp-Wave Ripples.2025. bioRxiv [Preprint]. 2025.07.24.666687.
- Khodagholy D, Gelinek JN, Buzsáki G. Learning-enhanced coupling between ripple oscillations in association cortices and hippocampus. *Science*. 2017.358(6361):369-372.
- Kveraga K, Ghuman AS, Bar M. Top-down predictions in the cognitive brain. *Brain Cogn*. 2007.65(2):145-68.
- Levashov, O. V., & Safiulina, V. F. (2026). The Role of Synchronization of Neural Modules in Pattern Processing in the Visual System. bioRxiv. <https://doi.org/10.64898/2025.12.29.696937>
- Malcolm GL, Henderson JM. The effects of target template specificity on visual search in real-world scenes: evidence from eye movements. *J Vis*. 2009. 9(11):8.1-13.
- Marr D. 1982. *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: W.H. Freeman.
- Mishra A, Akkol S, Espinal E, Markowitz N, Tostaeva G, Freund E, Mehta AD, Bickel S. Hippocampal sharp wave ripples and coincident cortical ripples orchestrate human semantic networks.2024.bioRxiv[Preprint]. 2024.04.10.588795.
- Okubo M, Nicholls ME. Hemispheric asymmetry in temporal resolution: contribution of the magnocellular pathway. *Psychon Bull Rev*. 2005.12(4):755-9.
- Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat Neurosci*. 2013.16(8):1068-76.
- Robinson HL, Todorova R, Nagy GA, Gruzdeva A, Paudel P, Oliva A, Fernandez-Ruiz A. Large sharp-wave ripples promote hippocampo-cortical memory reactivation and consolidation during sleep. *Neuron*. 2025.114(2):226-236.e6.
- Rochester N, Holland J, Haibt L and Duda W. "Tests on a cell assembly theory of the action of the brain, using a large digital computer," in *IRE Transactions on Information Theory*. 1956. vol. 2, no. 3, pp. 80-93.
- Trevelyan AJ, Sussillo D, Watson BO, Yuste R. Modular propagation of epileptiform activity: evidence for an inhibitory veto in neocortex. *J Neurosci*. 2006. 26(48):12447-55.
- Yang W, Sun C, Huszár R, Hainmueller T, Kiselev K, Buzsáki G. Selection of experience for memory by hippocampal sharp wave ripples. *Science*. 2024.383(6690):1478-1483.

Zeki S. A Vision of the Brain. Oxford : Blackwell Scientific Publications, 1993. 224 p

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.