Learning orientation detection system and its application to RGB images

Tianqi Chen[a], Zhiyu Qiu[a], Yuxiao Hua[a], Yuki Todo[*b], Zheng Tang[c]
[a]Division of Electrical Engineering and Computer Science, Graduate School of Natural Science & Technology, Kanazawa University, Kanazawa 920-1152, Ishikawa, Japan; [b]Faculty of Electrical, Information and Communication Engineering, Kanazawa University, Kanazawa, 920-1192, Ishikawa, Japan; [c]Faculty of Engineering, University of Toyama, Toyama 930-8555, Toyama, Japan

ABSTRACT

Our previous research showed that emulating dendritic neuron structure effectively addresses orientation detection challenges in learning tasks, reducing both learning time and costs compared to alternative neural network approaches. Our simulation model incorporates an On-Off Response mechanism in bipolar cells (BCs) and horizontal cells (HCs) for processing grayscale input images. To overcome limitations with single-channel input, we introduce color-selective cells. By integrating these cells, we enhance and select outputs of local orientation detection dendritic neurons, generating specific feature maps. These feature maps are generated using a biologically-inspired model that mimics the mechanism of color-selective cells. Additionally, global neurons are used to capture overall image features by aggregating outputs from local dendritic neurons. Our system utilizes backpropagation to update parameters of local orientation detection dendritic neurons. Furthermore, we integrate a learnable orientation detection neural network after the dendritic neuron stage.

Keywords: Artificial visual system, neural network, dendrite neuron, orientation detection, bio-inspired model

1. INTRODUCTION

Visual information, which includes aspects like shapes, colors, orientations and motion, plays a pivotal role in human behavior and decision making[1,2]. It guides our ability to recognize objects, navigate spaces, and identify faces, among other critical functions[3,4]. Orientation detection is particularly crucial as it underpins these abilities[5], orientation-selective cells respond to specific orientations[6,7]. However, even with significant progress in understanding the brain's orientation selectivity, the exact mechanisms by which the visual system detects orientations remain elusive. However, the exact mechanisms by which the visual system detects orientations remain elusive[8]. The development of neuron selectivity is believed to be influenced by external factors starting from birth, rather than being predetermined[9]. In response, researchers have looked to the dendritic model, a representation of biological neurons, to simulate the behavior of the human brain more accurately[10]. This model has been employed to address computer vision tasks, offering a simpler yet effective way to mimic the behavior of neurons in the primary visual cortex[11], also in our artificial visual system (AVS). It features receptive fields measuring 3*3, with each field specializing in a specific orientation. We also employ the On-Off Response mechanism, similar to that found in bipolar cells and horizontal cells[12]. This mechanism responds to light stimuli based on the relative brightness of the center compared to the surrounding areas. While horizontal cells exhibit center-surround receptive fields, BCs activate ganglion cells through on-center and off-center responses[13]. To extend our model's capabilities, we integrate a backpropagation-based learning algorithm, enhancing its learning ability. This allows us to process more complex images and derive weights and biases through training on a labeled image dataset. Integration of the dendritic neuron model enhances AVS accuracy, offering a biologically plausible simulation of orientation selectivity. Our model's use of an AVS algorithm, combined with an orientation detection neural network that makes signals processed through center-surround differences recognized by the cone cells[14], makes it capable of handling RGB input images and producing feature maps. By training on labeled image datasets, our AVS algorithm derives parameters, enabling feature detection from R, G, and B channels for each orientation. Computer simulations validate our model's efficacy in orientation detection and comparative analysis against traditional CNN-based systems. Notably, our system demonstrates resilience in scenarios where training data are scarce compared to CNN-based systems, resulting in lower learning costs. Additionally, our AVS exhibits superior robustness against salt and pepper

[*]yktodo@se.kanazawa-u.ac.jp

International Conference on Optics, Electronics, and Communication Engineering (OECE 2024), edited by Yang Yue, Proc. of SPIE Vol. 13395, 1339521 · © 2024 SPIE · 0277-786X · Published under a Creative Commons Attribution CC-BY 3.0 License · doi: 10.1117/12.3048289

noise due to local orientation detection dendritic neurons, less influenced by noise than CNNs capturing global features.

2. METHOD

2.1 Dendritic neuron model

The dendritic neuron model has been validated for its alignment with biological mechanisms, leveraging the functionality of dendrites and is crucial for computer visions[15,16]. To optimize the dendritic model for orientation detection, specific design considerations are crucial. Nonlinear processing mechanisms at the soma level can discern orientations using mathematical functions sensitive to orientation variations. Our adopted dendritic neuron model consists of four layers: synaptic, branches, membrane, and soma layers, each with distinct functions rooted in biology: Synaptic Layer: Represents synaptic connections between the dendritic neuron and other neurons; Branches Layer: Encompasses dendritic branches, providing a substantial surface area for input reception; Membrane Layer: Corresponds to the plasma membrane, regulating ion flux into and out of the cell; Soma Layer: Central component executing nonlinear processing of inputs. Figure 1 illustrates the simplified structure of a dendritic neuron. The diagram depicts a single soma connected to a system comprising *j* branches, denoted as B^m_1, B^m_2, ..., and B^m_i, forming the branches layer. The input signals x_1, x_2, ..., x_i originate from neighboring neurons and are linked to the thin orange lines representing branches. These inputs are connected to the branches via thick black lines, symbolizing synapses. The length of the thick black lines can be regarded as the relative distance parameter, d^m_{ij}, meaning the relative distance from the *i*th input neighboring neuron to the *j*th branch of the *m*th soma. Synapses are indicated by either exhibitory connections or inhibitory connections. The functions attributed to the various layers of a dendritic neuron model collectively contribute to the creation of intricate processing mechanisms within the neuron. Based on this function, the specific functions of each layer can be customized to suit the requirements of the applications of the model. In this study, we utilize the sigmoid function as the activation function within synapses, thereby facilitating synaptic learning processes. The sigmoid function plays a crucial role in determining the connection strength between neighboring input neurons and dendritic branches, which is influenced by the weight and bias of the synapse. This approach enables the dendritic neuron model to adaptively adjust its synaptic connections based on input signals, enhancing its capability to learn and process information effectively. Due to the constraints of the neural signal range, only the interval [0,1] will be considered. Several scenarios should be considered due to the intricate relationship among connection positions and the interplay among *w*, *q*, and 0, presented as 4 connection positions: constant-0, constant-1, exhibitory and inhibitory. Adjusting weights and biases through learning algorithms shapes connectivity among dendritic neurons, akin to human learning and memory. This dynamic modification enables the formation, pruning, and adaptation of connections, reflecting human brain processes[17,18]. Through learning, the relevant connections are reinforced. The dynamic adjustment shows the plasticity of model, allowing the model to learn from experience and enhance performance over time.

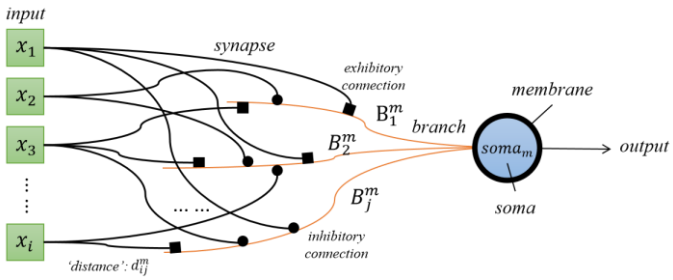

Figure 1. A sample of a part of dendritic neuron model.

2.2 Local orientation detection dendrite neuron model & on-off response mechanism

In our model, the retina processes visual information and detects object orientations. The receptive field sized 3*3 includes a central pixel and its 8 surrounding pixels. Each local orientation detection dendritic neuron receives inputs from 9 pixels and discerns orientations based on active signals in 2 or 3 pixels, corresponding to 0, 45, 90, and 135 degrees. Given the various orientations, we allocate 4 somas, O_1 to O_4, for detecting orientations. Signals from central and surrounding pixels, ranging from 0 to 1, are used. For full-color images with 3 input channels (scale 0-255), we employ the On-Off Response mechanism in BCs and HCs. In the visual system, light stimuli trigger responses in receptive fields. When the central area receives more light (on response), HCs inhibit adjacent photoreceptors, enhancing contrast. Conversely, decreased light (off response) reduces this inhibition, allowing signal transmission to BCs, and

further enhancing contrast[19]. Ganglion cells with on-off responses detect local contrast and edges, crucial for visual scene analysis. They integrate signals from diverse BCs, aiding in contrast sensitivity and encoding visual data for cortical processing. This mechanism in ganglion cells facilitates color differentiation in the cortex. BCs amplify signals from photoreceptors, aiding in stimulus discrimination[20,21]. BCs amplify signals from photoreceptors, aiding in stimulus discrimination. Image scanning perceives scale value variations as light stimulus differences. Photoreceptors convert stimuli into signals for HCs and BCs. BC simulation neurons mimic BCs' function, transforming light stimuli into signals. Inputs from the same receptive field may differ post-recognition. HCs compare differences to a threshold, inducing excitatory or inhibitory responses. Outputs of HC and BC simulation neurons illustrate disparities between central and neighboring photoreceptor signals. Other outputs represent regularized pixel scale values (0-1). After processing by HC and BC simulation neurons, parameters are normalized for dendritic neuron calculation, the sample of the HC and BC preprocessing is shown in Figure 2, where the '?' in the circle means an unsure connection position before learning. Learning prunes irrelevant synapses/branches, mimicking vertebrate feature memorization. Figure 3 refines the local orientation detection dendritic neuron model, leaving essential components. Soma responds only to relevant inputs, exhibiting excitatory or inhibitory responses. Learning yields four specialized orientation detection neurons, aligning with Hubel and Wiesel's theory. With a nine-pixel receptive field, our model achieves local orientation detection, mirroring cortex neuron design for orientation selectivity in detection tasks.

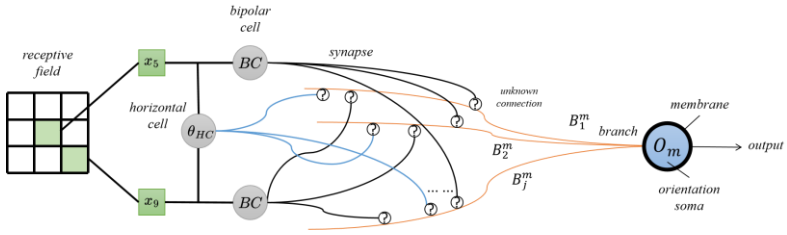

Figure 2. The structure of On-Off Response in the model.

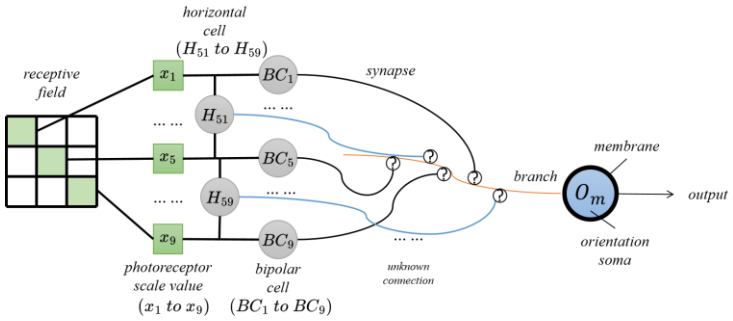

Figure 3. The structure of a local orientation detection dendritic neuron in the model. All the pixels in the photoreceptor have their own bipolar cells simulation neuron and horizontal cells simulation neuron.

2.3 Global orientation detection dendrite neuron model & feature map generation

For an M×N pixel image, the local orientation detection neuron remains effective. The global orientation detection dendritic neuron model is built upon analyses from multiple receptive fields by the local orientation detection dendritic neuron model. In this study, we opt to detect the image's global orientation by scanning all pixels, with each pixel serving as a receptive field center covering the entire image. Thus, each of the four local orientation detection dendritic neurons produces M×N outputs corresponding to a specific orientation, indicating activations of local motion directive detection neurons. Post-learning, excitatory reactions may occur in receptive fields centered on pixels receiving light stimuli. To manage outputs from the learned global orientation detection dendritic neuron model, specific cell calculations are necessary. Retinal research suggests global detection is achieved by summing local neuron outputs and comparing differences, primarily within the primary cortex. The need for a sum pooling layer after local edge-orientation detection neurons is crucial to determine the correct angle in the presence of multiple detection results. This layer aggregates inputs to provide a summarized output, allowing our model to evaluate somas' activations for each orientation and determine the final output of each channel. In 3-channel images, consistent orientation may not be observed across all channels, demanding cone cells' cooperative behavior akin to an AND gate relationship. Cone cells, pivotal for color selectivity in the retina, process color information essential for object recognition and perception. Our

AVS integrates cone cell simulation neurons for each orientation, simplifying subsequent processing with summing operations and aligning with biological brain models. Figure 4 illustrates the complete diagram of the global orientation detection dendritic neuron in our AVS. This dendritic neuron evaluates information from local orientation detection neurons across successive receptive fields, covering every pixel in the image, and gathers orientation information for the entire image. Activated soma signals are normalized by cone cell simulation neurons to construct a one-channel feature map for each orientation. Ultimately, the orientation angle with the highest number of activated somas is outputted, enabling efficient orientation detection across the entire image.

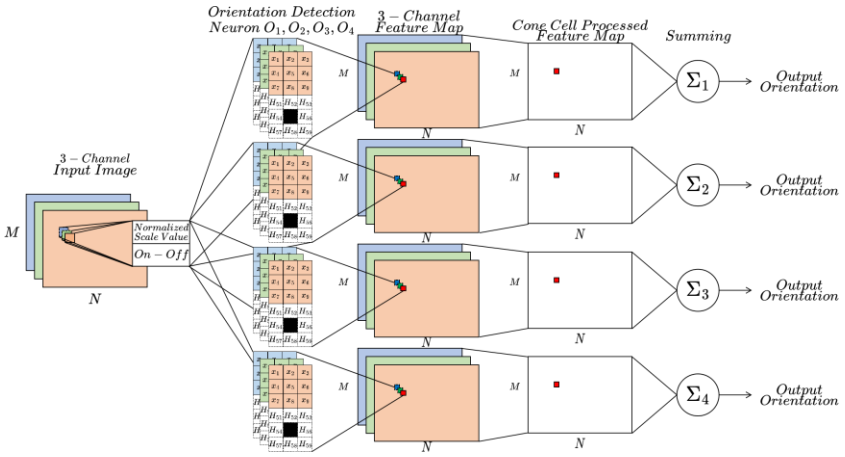

Figure 4. The diagram of the orientation detection system on a single 3-channel image.

3. SIMULATION RESULT

In our experimental simulations, CNNs are prominent for orientation detection due to their efficiency and accuracy, excelling in learning complex features from large datasets, even with high-dimensional inputs. Additionally, CNN architectures exhibit remarkable generalization and robustness across various scenarios. However, CNNs require substantial training data and computational resources, limiting scalability. In contrast, our model, with just 144 parameters, offers simplicity and adaptability, leveraging dendritic neurons. So, we compared our model with CNN-based orientation detection systems, including 4-layer CNNs, LeNet, EfficientNet-B0 (pretrained or not), and ResNet50. Full-color images are crucial in many applications, capturing a wide spectrum of light stimuli perceivable by human eyes. Our study utilized a dataset of 128×128 3-channel images, as shown in Figure 5a. The dataset comprises meticulously crafted images, ranging from 3 to over 256 pixels in object size, with consistent scale values within each object or randomly generated values for clarity. Randomly generated scale values are also included for the background, enhancing dataset diversity. This variety in object size allows for a comprehensive evaluation of the model's performance across different scales and ensures robustness and generalization capabilities. The dataset contains 20,000 images, evenly distributed across four orientations, with 10,000 images used for training. All models employ identical hyperparameters to ensure fair comparison. Initially, we examine the influence of the training data ratio. Traditional neural networks, typically require large datasets for effective training and aim to capture global features corresponding to labels. Insufficient data can hinder the ability of these networks to extract meaningful information. In contrast, our model leverages local orientation detection of dendritic neurons, enabling it to capture information effectively from single images containing numerous features relevant to the corresponding orientation. To substantiate this claim, we maintain the dataset composition but vary the ratio of training data. Specifically, we explore data ratios of 75:25, 50:50, 10:90, and 5:95. The stability of the model is crucial in computer vision, particularly when the number of training data decreases significantly. The simulation results are presented in Table 1. Simulation results show that while all CNNs excel in recognizing features and detecting orientations, simpler architectures like LeNet and 4-Layer CNN exhibit lower stability in test accuracy with low training data ratios, prone to overfitting. In contrast, complex structures like EfficientNet-B0 (EfN-un) and ResNet50 demonstrate greater stability, less influenced by variations in training data ratios. Pretrained EfficientNet-B0 (EfN-B0) is the least sensitive to training data ratios due to its training on ImageNet. However, our AVS, trained on the same dataset, may not achieve the highest accuracies with ample training data but demonstrates superior stability with decreasing training data ratios, thanks to its effective orientation information extraction from individual images. To assess noise immunity, we employed Salt and Pepper noise, simulating real-world disruptions and

potential overfitting, recording Test Acc in Table 2. These images underwent rotation for varied orientations and meticulous labeling to represent human-detected orientation, as shown in Figure 5b. Object sizes were randomly determined for dataset variability. This real-world dataset enriches model training and testing diversity. The models from Table 1 were used for processing. Our AVS demonstrates excellent noise immunity compared to other models, particularly in handling Salt & Pepper noise. Pretrained EfficientNet-B0 tends to overfit, capturing excessive unnecessary information from our training dataset. Other models exhibit lower noise immunity with real-world object datasets or fail to capture real object orientation information. Nonetheless, AVS consistently outperforms other models.

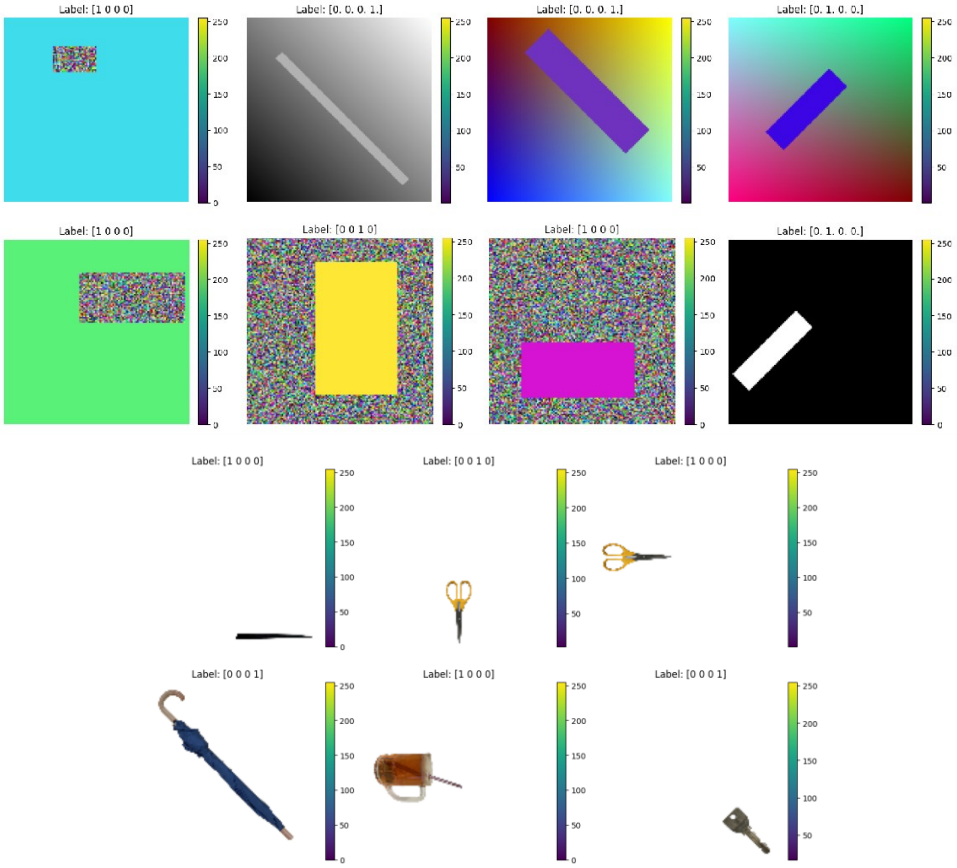

Figure 5. The example of (a) the images in our dataset and corresponding labels and (b) the real object images.

Table 1. Accuracy of artificial neural network models facing different training data ratios.

Model	Train: Test							
	75:25		50:50		10:90		5:95	
Dendrite	Train Acc	Test Acc	Train Acc	Test Acc	Train Acc	Test Acc	Train Acc	Test Acc
AVS	99.53±0.05%	99.46±0.57%	99.28±0.35%	99.21±0.48%	99.21±2.06%	98.61±1.99%	98.43±1.07%	95.04±4.30%
LeNet	98.26±1.06%	95.16±1.81%	99.16±0.66%	94.52±0.90%	98.59±0.49%	79.8±4.21%	98.38±0.79%	63.95±4.15%
4LCNN	99.94±0.11%	98.50±0.44%	100.00%	97.36±1.06%	100.00%	80.22±2.21%	100.00%	61.72±3.42%
EfN-B0	99.99±0.00%	99.98±0.03%	100.00%	99.95±0.06%	100.00%	99.87±0.07%	100.00%	99.8±0.08%
EfN-un	99.96±0.00%	99.92±0.04%	99.99±0.01%	99.69±0.25%	99.99±0.01%	92.99±1.00%	99.97±0.03%	81.12±2.17%
ResNet50	99.86±0.11%	99.77±0.24%	99.99±0.01%	99.84±0.08%	99.87±0.41%	95.56±0.30%	99.95±0.16%	79.84±4.88%

Table 2. Accuracy of artificial neural network models facing different numbers of salt&pepper noise and real objects.

Data ratio						
Model	**75:25**			**50:50**		
Data	**5%**	**10%**	**Object**	**5%**	**10%**	**Object**
AVS	98.78±5.56%	73.83±10.68%	95.31±5.06%	98.83±5.12%	70.45±12.1%	95.03±2.71%
LeNet	60.85±3.14%	36.48±2.44%	86.13±1.94%	58.91±6.75%	35.02±3.03%	83.37±2.21%
4LCNN	60.01±6.72%	40.00±2.90%	91.81±1.31%	58.89±5.28%	38.12±2.64%	88.68±2.44%
EfN-B0	86.19±4.81%	57.88±6.34%	98.47±1.05%	87.84±5.84%	57.59±8.59%	98.65±1.02%
EfN-un	78.50±7.01%	63.72±3.74%	98.27±0.37%	79.75±4.04%	63.19±5.34%	97.16±1.13%
ResNet50	87.55±5.89%	80.15±6.47%	98.16±0.42%	86.52±6.67%	77.11±6.09%	95.48±3.22%
Model	**10:90**			**5:95**		
Data	**5%**	**5%**	**10%**	**5%**	**10%**	**Object**
AVS	97.93±7.38%	97.46±4.94%	61.44±15.1%	97.46±4.94%	61.44±15.1%	90.92±1.12%
LeNet	40.31±9.09%	34.75±2.22%	28.80±1.18%	34.75±2.22%	28.80±1.18%	53.49±3.77%
4LCNN	48.54±3.49%	38.42±2.82%	29.79±1.30%	38.42±2.82%	29.79±1.30%	46.48±3.52%
EfN-B0	94.27±2.57%	96.45±1.91%	82.97±4.46%	96.45±1.91%	82.97±4.46%	96.78±4.28%
EfN-un	62.37±5.66%	33.83±3.08%	26.70±0.54%	33.83±3.08%	26.70±0.54%	69.19±4.78%
ResNet50	83.39±6.93%	58.43±10.1%	39.96±5.89%	58.43±10.1%	39.96±5.89%	61.67±10.9%

4. CONCLUSIONS

In conclusion, our AVS represents a significant advancement in computer vision, drawing inspiration from the human visual system's biological mechanisms. By closely simulating components of the retina such as dendritic structures, cone cells, and the On-Off Response mechanism, our model demonstrates a high degree of biological plausibility. One key strength lies in its adaptiveness, particularly with limited training data. Unlike traditional CNN models, our bio-inspired approach enables efficient learning with smaller datasets, enhancing practicality in real-world scenarios. Furthermore, our AVS exhibits higher robustness compared to CNN models, maintaining stable performance amidst noise or data variations. Importantly, our model bridges the gap between brain science and computer vision by mimicking biological processes like dendritic neuron formation, offering insights into the human visual system's workings. This interdisciplinary approach holds promise for advancing neuroscience and artificial intelligence.

ACKNOWLEDGMENTS

This work was supported by JST SPRING, Grant Number JPMJSP2135, and JSPS KAKENHI, Grant Number 23K11261.

REFERENCES

[1] Todo, Y., Tang, Z., Todo, H., Ji, J. and Yamashita, K., "Neurons with multiplicative interactions of nonlinear synapses," International Journal of Neural Systems, 29(08), 1950012 (2019).

[2] Medina, J., [Brain Rules: 12 Principles for Surviving and Thriving at Work, Home, and School], ReadHowYouWant.com, 2011.
[3] Vanston, J. E. and Strother, L., "Sex differences in the human visual system," Journal of Neuroscience Research, 95(1-2), 617-625 (2017).
[4] Namboodiri, V. M. K., Huertas, M. A., Monk, K. J., Shouval, H. Z. and Shuler, M. G. H., "Visually cued action timing in the primary visual cortex," Neuron, 86(1), 319-330 (2015).
[5] Ferster, D. and Miller, K. D., "Neural mechanisms of orientation selectivity in the visual cortex," Annual Review of Neuroscience, 23(1), 441-471 (2000).
[6] Hubel, D. H. and Wiesel, T. N., "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," The Journal of Physiology, 160(1), 106 (1962).
[7] Hubel, D. H., "Exploration of the primary visual cortex, 1955-78," Nature, 299(5883), 515-524 (1982).
[8] Hubel, D. H. and Wiesel, T. N., "Receptive fields and functional architecture of monkey striate cortex," The Journal of Physiology, 195(1), 215-243 (1968).
[9] Blakemore, C. and Cooper, G. F., "Development of the brain depends on the visual environment," Nature, 228(5270), 477-478 (1970).
[10] Ji, J., Tang, C., Zhao, J., Tang, Z. and Todo, Y., "A survey on dendritic neuron model: Mechanisms, algorithms and practical applications," Neurocomputing, 489, 390-406 (2022).
[11] Hua, Y., Todo, Y., Tang, Z., Tao, S., Li, B. and Inoue, R., "A novel bio-inspired motion direction detection mechanism in binary and grayscale background," Mathematics, 10(20), 3767 (2022).
[12] Turner, M. H., Schwartz, G. W. and Rieke, F., "Receptive field center-surround interactions mediate context-dependent spatial contrast encoding in the retina," Elife, 7, e38841 (2018).
[13] Schiller, P. H., Sandell, J. H. and Maunsell, J. H. R., "Functions of the on and off channels of the visual system," Nature, 322(6082), 824-825 (1986).
[14] Derrington, A. M., Krauskopf, J. and Lennie, P., "Chromatic mechanisms in lateral geniculate nucleus of macaque," The Journal of Physiology, 357(1), 241-265 (1984).
[15] Vlasits, A. and Baden, T., "Motion vision: A new mechanism in the mammalian retina," Current Biology, 29(19), R933-R935 (2019).
[16] Rossi, L. F., Harris, K. D. and Carandini, M., "Spatial connectivity matches direction selectivity in visual cortex," Nature, 588(7839), 648-652 (2020).
[17] Kandel, E. R., "The molecular biology of memory storage: a dialogue between genes and synapses," Science, 294(5544), 1030-1038 (2001).
[18] Squire, L. R. and Kandel, E. R., "Memory: From Mind to Molecules," Scientific American Library (1999).
[19] Dacheux, R. F. and Raviola, E., "The rod pathway in the rabbit retina: a depolarizing bipolar and amacrine cell," Journal of Neuroscience, 6(2), 331-345 (1986).
[20] Hanazawa, A., Komatsu, H. and Murakami, I., "Neural selectivity for hue and saturation of color in the primary visual cortex of the monkey," European Journal of Neuroscience, 12(5), 1753-1763 (2000).
[21] Lennie, P., Krauskopf, J. and Sclar, G., "Chromatic mechanisms in striate cortex of macaque," Journal of Neuroscience, 10(2), 649-669 (1990).