Biological Cybernetics (2023) 117:297–298
https://doi.org/10.1007/s00422-023-00977-6

Biological Cybernetics

EDITORIAL

What can computer vision learn from visual neuroscience? Introduction to the special issue

Kexin Chen[1] · Hirak J. Kashyap[2] · Jeffrey L. Krichmar[3] · Xiumin Li[4]

Published online: 9 October 2023

This special issue on "What Can Computer Vision Learn From Visual Neuroscience?" follows a workshop with the same name organized by the editors at the Computer Vision and Pattern Recognition (CVPR) 2022 conference, where researchers from all over the world presented their ideas and findings on integrating computational theories in the brain into machine vision models and systems. The overarching goal of the workshop and the presented research was to make progress toward developing efficient and robust artificial vision systems, like the ones possessed by humans and other biological organisms. Given the broad nature of the research question posed in the workshop, lectures and original contributions discussed a variety of bio-inspired applications to machine vision, such as hierarchical representation of visual concepts, event-based asynchronous computations and spiking neural networks, active vision for robotic systems and robust image recognition, among others.

In the last decade, machine vision has seen tremendous progress using deep learning for both high-level and low-level tasks, such as object recognition, segmentation, motion and depth perception. However, machine vision still lags behind human vision in many real-world scenarios, where the data distributions are unknown, evolving and long tailed. The current approaches in computer vision also deviate from characteristics of biological vision systems that make them efficient. For example, in biological vision, learning occurs through continuous interactions of the biological agent (i.e., a body) with the world. On the other hand, the learning process in machine vision is usually disconnected from the environment. Also, the learning-based machine vision models are easily deceived by adversarial inputs. However, the same adversarial examples are unmistakably recognized by humans, implying that the feature representations learned by the current deep learning models are not robust and conceptual. Furthermore, the brain operates using a small amount of energy, unlike the state-of-the-art vision systems that run on graphics processors that consume many kilowatts of power. Although brain-inspired neuromorphic vision sensors and processors could potentially lead to low-power machine vision systems, they are yet to be scaled up algorithmically for large-scale real-world applications. New inspirations from the brain could hence make machine vision more efficient, robust and capable of continuous learning and adaptation. This goal also requires understanding of the brain's visual processing and learning mechanisms for diverse sets of tasks and incorporating them into the corresponding computational models.

This special issue brings together contributions from some of the workshop participants and additional researchers on how ideas learned from visual neuroscience can improve computer vision systems. They address many of the issues described above and show how ideas from neuroscience can be incorporated into computer vision.

The functional neuroanatomy of the visual system is structurally different from most artificial neural networks (ANNs). Schmid et al. discuss canonical principles of brain computations and provide formal specifications for the principles that utilize recurrent feedforward, lateral and feedback interactions found in the neuroscience literature; the authors show how they can be utilized to define mechanisms for visual shape and motion processing (Schmid et al. 2023).

Despite the superhuman performance in deep learning on image classification, humans and ANNs process visual information quite differently. Malik et al. address the problem of adversarial robustness for computer vision and investigate the type of image transformations that humans are robust to and machine learning systems are not (Malik et al. 2023).

✉ Jeffrey L. Krichmar
jkrichma@uci.edu

1 Department of Cognitive Sciences, University of California, Irvine, USA

2 Department of Computer Science, University of California, Irvine, USA

3 Department of Cognitive Sciences, Department of Computer Science, University of California, Irvine, USA

4 College of Automation, Chongqing University, Chongqing, China

These transformations could guide training of future neural networks to improve robustness to adversarial attacks.

Sparse coding is ubiquitous in the brain, especially in the visual system, and may be a canonical feature of nervous systems. Lässig et al. hypothesize that error backpropagation-based training of neural networks leads to catastrophic forgetting due to its tendency to generate distributed and overlapping representations (Lässig et al. 2023).

In addition to image classification and motion detection, the visual nervous system must rapidly perceive and make the appropriate decisions based on incoming visual information. Briden and Norouzi apply deep neural network techniques to analyze electroencephalograms (EEG) and predict human perceptual decision-making confidence (Briden & Norouzi 2023). Their method also identifies the brain regions involved in the task, thus guiding visual neuroscience research using ANNs.

Unlike ANNs, nervous systems have delays between neurons due to axon and dendritic tree properties. Grimaldi and Perrinet take inspiration from varied synaptic delays in dendritic trees to develop a visual motion detection model that is robust to pruning and can be used by neuromorphic applications to reduce computations (Grimaldi & Perrinet 2023). They propose a solution with winner-take-all sparsity and within-layer connections, grounded in neuroscience, to learn sparse and non-overlapping representations that help to reduce catastrophic forgetting.

Neuromorphic vision sensors operate more like the eye than regular cameras in that the sensor only responds to pixel intensity changes rather than processing a complete frame every so many milliseconds. Gruel et al. show that foveation, which occurs in the mammalian vision system, can be successfully applied to neuromorphic event camera data to perform scene understanding at high resolution and with significantly reduced computations, while not sacrificing accuracy compared to full-frame processing commonly used in computer vision (Gruel et al. 2023).

We hope this special issue on "What Can Computer Vision Learn From Visual Neuroscience?" can work in both directions: Inspiring computer scientists and artificial intelligence enthusiasts to incorporate more neuroscience into their systems, and encouraging computational neuroscientists and visual neuroscientists to improve their knowledge by using these cutting edge methods.

References

Briden M, Norouzi N (2023) Toward metacognition: subject-aware contrastive deep fusion representation learning for EEG analysis. Biol Cybern. https://doi.org/10.1007/s00422-023-00967-8

Grimaldi A, Perrinet LU (2023) Learning heterogeneous delays in a layer of spiking neurons for fast motion detection. Biol Cybern. https://doi.org/10.1007/s00422-023-00975-8

Gruel A, Hareb D, Grimaldi A, Martinet J, Perrinet L, Linares-Barranco B, Serrano-Gotarredona T (2023) Stakes of neuromorphic foveation: a promising future for embedded event cameras. Biol Cybern. https://doi.org/10.1007/s00422-023-00974-9

Lässig F, Aceituno PV, Sorbaro M, Grewe BF (2023) Bio-inspired, task-free continual learning through activity regularization. Biol Cybern. https://doi.org/10.1007/s00422-023-00973-w

Malik G, Crowder D, Mingolla E (2023) Extreme image transformations affect humans and machines differently. Biol Cybern. https://doi.org/10.1007/s00422-023-00968-7

Schmid D, Jarvers C, Neumann H (2023) Canonical circuit computations for computer vision. Biol Cybern. https://doi.org/10.1007/s00422-023-00966-9