curvature to induce stiffening of this arch to an extent similar to that of modern humans. For example, the authors examined the species *Australopithecus afarensis*. This species existed more than three million years ago, and whether it walked upright in a human-like fashion is debated[6–8]. Venkadesan *et al.* report that the transverse arch of *A. afarensis* was less curved than that of a human foot and thus, according to their model, probably less stiff. However, the authors correctly emphasize that such curvature alone cannot be used reliably to infer movement capabilities, and other mechanisms might stiffen the foot sufficiently to allow a human-like gait.

The curvature of transverse arches in human populations probably spans a wide range of values. Some people have noticeably flat feet whereas others have a high arch. Perhaps those with flat feet have less curvature of their transverse arch and thus potentially reduced stiffness in their feet compared with those whose feet are less flat. But it is also possible that people with flat feet have sufficient transverse-arch curvature to compensate for their low longitudinal arch, thereby maintaining sufficient stiffness for effective walking and running. Given that Venkadesan and colleagues' work did not directly test whether there is a relationship between transverse-arch curvature and the stiffness of the human foot, it remains to be determined whether the range of differences in human transverse-arch curvature is a crucial functional parameter to explain foot stiffness.

The range of curvature of the arch of human feet suggested by Venkadesan *et al.* would indicate that a nearly twofold change in stiffness is possible as a result of natural variation in curvature of the trans verse arch from one person to the next. However, any relationship between transverse-arch curvature and stiffness is probably not enough to completely explain the regulation of foot stiffness, and other factors will also need to be considered — for example, the stiffness of the plantar fascia or the potential for muscles to actively regulate arch stiffness. As such, caution is necessary before relying on this curvature parameter alone as the key variable in assessing human foot stiffness.

The fields of evolutionary biology, sports science and medicine have largely neglected the transverse arch when trying to explain the managements of loads applied to the foot. Venkadesan and colleagues' research suggests a new mechanism that links foot form and function and sets the scene for a possible shift in how the human foot is considered. More research will be needed to better understand how the transverse arch contributes to human locomotor performance, including determining what its contribution is to an individual's foot stiffness and whether this provides any mechanical or energetic benefits. It is conceivable that new treatments that take advantage of transverse-arch curvature to modulate foot stiffness could be developed for various foot disorders. Perhaps even more exciting are the implications of this work for efforts to mimic a human foot when designing prosthetic limbs or legged robots.

**Glen A. Lichtwark** and **Luke A. Kelly** are at the University of Queensland, School of Human Movement and Nutrition Sciences, St Lucia, Queensland 4072, Australia.
e-mails: g.lichtwark@uq.edu.au; l.kelly3@uq.edu.au

1. Venkadesan, M. *et al. Nature* **579**, 97–100 (2020).
2. Pontzer, H. *Curr. Biol.* **27**, R613–R621 (2017).
3. Holowka, N. B. & Lieberman, D. E. *J. Exp. Biol.* **221**, eb174425 (2018).
4. Ker, R. F., Bennett, M. B., Bibby, S. R., Kester, R. C. & Alexander, R. M. *Nature* **325**, 147–149 (1987).
5. Ward, C. V., Kimbel, W. H. & Johanson, D. C. *Science* **331**, 750–753 (2011).
6. Jungers, W. L. *Nature* **297**, 676–678 (1982).
7. Ward, C. V. *Am. J. Phys. Anthropol.* **119** (Suppl. 35), 185–215 (2002).
8. Hatala, K. G., Demes, B. & Richmond, B. G. *Proc. R. Soc. B* **283**, 20160235 (2016).

**The authors declare competing financial interests: see go.nature.com/2vefydy for details.**

**This article was published online on 26 February 2020.**

Artificial intelligence

# In-sensor computing for machine vision

**Yang Chai**

An image-sensor array has been developed that acts as its own artificial neural network to capture and identify optical images simultaneously, processing the information rapidly without needing to convert it to a digital format. **See p.62**

Sight is one of our most vital senses. Biologically inspired machine vision has developed rapidly in the past decade, to the point that artificial systems can 'see' in the sense of gaining valuable information from images and videos[1,2], although human vision remains much more efficient. On page 62, Mennel *et al.*[3] report a design for a visual system that, rather like the brain, can be trained to classify simple images in nanoseconds.

Modern image sensors such as those in digital cameras are based on semiconductor (solid-state) technology and were developed in the early 1970s; they fall into two main types, known as charge-coupled devices and active-pixel sensors[4]. These sensors can faithfully capture visual information from the environment, but generate a lot of redundant data. This vast amount of optical information is usually converted to a digital electronic format and passed to a computing unit for image processing.

The resulting movement of massive amounts of data between sensor and processing unit results in delays (latency) and high power consumption. As imaging rates and numbers of pixels grow, bandwidth limitations make it difficult to send everything back to a centralized or cloud-based computer rapidly enough for real-time processing and decision-making — which is especially important for delay-sensitive applications such as driverless vehicles, robotics or industrial manufacturing.

A better solution would be to shift some of the computational tasks to the sensory devices at the outer edges of the computer system, reducing unnecessary data movement. And because sensors normally produce analog (continuously varying) outputs, analog processing would be preferable to digital: analog-to-digital conversion is notoriously time- and energy-consuming.

To mimic the brain's efficient processing of information, biologically inspired neuromorphic engineering adopts a computing architecture that has highly interconnected elements (neurons, connected by synapses), allowing parallel computing (Fig. 1a). These artificial neural networks can learn from their surroundings by iteration — for instance, learning to classify something after being shown known examples (supervised learning), or to recognize a characteristic structure of an object from input data without extra information (unsupervised learning). During learning, an algorithm repeatedly makes predictions and strengthens or weakens each synapse in the network until it reaches an optimum setting.

Mennel and co-workers implement an artificial neural network directly in their image sensor. On a chip, they construct a network of photodiodes — tiny, light-sensitive units, each consisting of a few atomic layers of tungsten diselenide. This semiconductor's response to light can be increased or decreased by altering an applied voltage, so that the sensitivity of each diode can be individually tuned. In effect, this turns the photosensor network into a neural network (Fig. 1b) and allows it to carry out simple computational tasks. Changing the light responsivity of a photodiode alters the connection strength — the synaptic weight — in the network. Thus, the device combines optical sensing with neuromorphic computing.

The authors arrange the photodiodes into a square array of nine pixels, with three diodes to each pixel. When an image is projected on to the chip, various diode currents are produced, combined and read. The hardware array provides a form of analog computing: each photodiode generates an output current that is proportional to the incident light intensity, and the resulting currents are summed along a row or column, according to Kirchhoff's law (a fundamental rule of currents in circuits).

The array is then trained to perform a task. The discrepancy between the currents produced by the array and the predicted currents (the currents that would be produced if the array responds correctly to the image, for a given task) is analysed off-chip and used to adjust the synaptic weight for the next training cycle. This learning stage takes up time and computing resources, but, once trained, the chip performs its set task rapidly.

Using different algorithms for the neural network, the authors demonstrate two neuromorphic functions. The first is classification: their $3 \times 3$ array of pixels can sort an image into one of three classes that correspond to three simplified letters, and thus identify which letter it is in nanoseconds. This relatively simple task is just a proof of concept, and could be extended to recognizing more-complicated images if the array size were scaled up.

The second function is autoencoding: the computing-in-sensor array can produce a simplified representation of a processed image by learning its key features, even in the presence of signal noise. The encoded version contains only the most essential information, but can be decoded to reconstruct an image close to the original.

There is more to be done before this promising technology can be used in practical applications. A neuromorphic visual system for autonomous vehicles and robotics will need to capture dynamic images and videos in three dimensions and with a wide field of view. Currently used image-capture technology usually translates the 3D real world into 2D information, thereby losing movement information and depth. The planar shape of existing image-sensor arrays also restricts the development of wide-field cameras[5].

Imaging under dim light would be difficult for the device described by the authors. A redesign would be needed to improve light absorption in the thin semiconductor and to increase the range of light intensities that can be detected. Furthermore, the reported design requires high voltages and consumes a lot of power; by comparison, the energy consumption per operation in a biological neural network is at the sub-femtojoule level ($10^{-15}$ to $10^{-13}$ joules)[6]. It would also be useful to expand the response to ultraviolet and infrared light, to capture information unavailable in the visible spectrum[7].

The thin semiconductors used are difficult to produce uniformly over large areas, and are hard to process so that they can be integrated with silicon electronics, such as external circuits used for readout or feedback control. The speed and energy efficiency of devices that use these sensors will be dominated not by the image-capturing process, but by data movement between sensors and external circuits. Moreover, although the computing-in-sensor unit collects and computes data in the analog domain, reducing analog-to-digital conversions, the peripheral circuits still suffer from other intrinsic delays. The sensors and external circuits will need to be co-developed to decrease the latency of the entire system.

Mennel and colleagues' computing-in-sensor system should inspire further research into artificial-intelligence (AI) hardware. A few companies have developed AI vision chips based on silicon electronics[8], but the chips' intrinsic digital architecture leads to problems of latency and power efficiency.

More broadly, the authors' strategy is not limited to visual systems. It could be extended to other physical inputs for auditory, tactile, thermal or olfactory sensing[9–11]. Development of such intelligent systems, together with the arrival of the 5G fast wireless network, should allow real-time edge (low-latency) computing in the future.

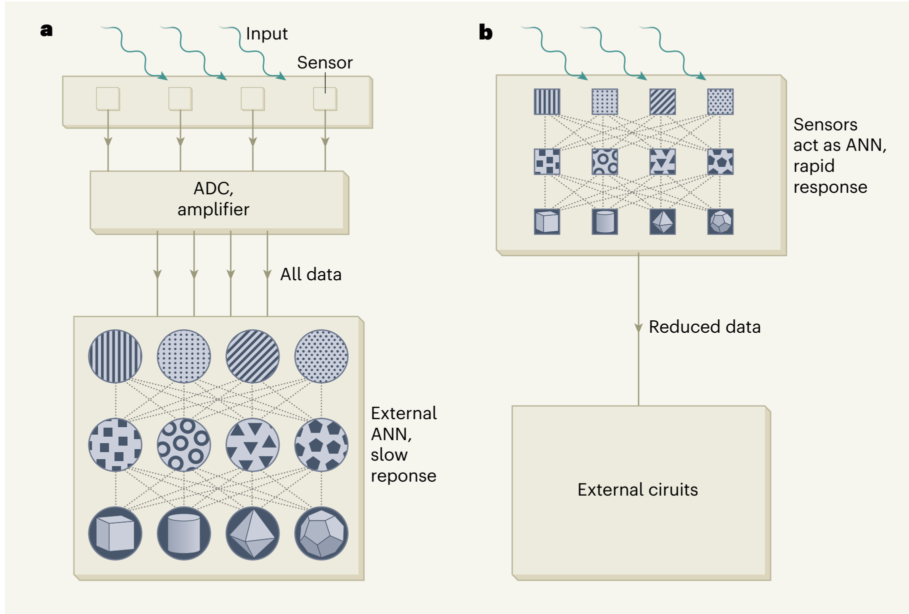


**Figure 1 | Computing within a vision sensor for intelligent and efficient preprocessing. a**, In conventional artificial-intelligence (AI) vision sensors, signals are collected from light-responsive sensors, converted from analog to digital form (ADC, analog-to-digital converter), amplified and then fed as inputs to an external artificial neural network (ANN) — layers of interconnected computational units (circles) whose connections can be adjusted, allowing the network to be trained to perform tasks such as classifying images. An input layer of the ANN receives signals encoding simple physical elements (represented here by dots and lines); in subsequent layers, these are optimized to mid-level features (simple shapes); and refined images are formed at the output layer (3D shapes). The overall response can be slow and energy-hungry. **b**, Mennel *et al.*[3] report a system in which interconnected sensors (squares) on a chip not only collect signals, but also work as an ANN to recognize simple features, reducing movement of redundant data between sensors and external circuits.

**Yang Chai** is in the Department of Applied Physics, Hong Kong Polytechnic University, Kowloon, Hong Kong, P. R. China.
e-mail: ychai@polyu.edu.hk

1. Mead, C. *Proc. IEEE* **78**, 1629–1636 (1990).
2. Kyuma, K. *et al. Nature* **372**, 197–198 (1994).
3. Mennel, L. *et al. Nature* **579**, 62–66 (2020).
4. Bigas, M., Cabruja, E., Forest, J. & Salvi, J. *Microelectr. J.* **37**, 433–451 (2006).
5. Choi, C. *et al. Nature Commun.* **8**, 1664 (2017).
6. Laughlin, S. B., de Ruyter van Steveninck, R. R. & Anderson, J. C. *Nature Neurosci.* **1**, 36–41 (1998).
7. Zhang, K. *et al. ACS Nano* **10**, 3852–3858 (2016).
8. Davies, M. *et al. IEEE Micro* **38**, 82–99 (2018).
9. Wan, C. *et al. Adv. Mater.* **30**, 1801291 (2018).
10. Kim, Y. *et al. Science* **360**, 998–1003 (2018).
11. Qu, T. Y. *et al. Adv. Mater.* **32**, 1907288 (2020).