Artificial neural networks enabled by nanophotonics

Qiming Zhang¹, Haoyi Yu¹, Martina Barbiero¹, Baokai Wang¹ and Min Gu¹

Abstract
The growing demands of brain science and artificial intelligence create an urgent need for the development of artificial neural networks (ANNs) that can mimic the structural, functional and biological features of human neural networks. Nanophotonics, which is the study of the behaviour of light and the light–matter interaction at the nanometre scale, has unveiled new phenomena and led to new applications beyond the diffraction limit of light. These emerging nanophotonic devices have enabled scientists to develop paradigm shifts of research into ANNs. In the present review, we summarise the recent progress in nanophotonics for emulating the structural, functional and biological features of ANNs, directly or indirectly.

Introduction
The human brain is the most complex biological organ in the universe and remains unknown to us. Several countries have started large programmes or alliances on brain science¹. Together, these initiative projects will undoubtedly lead to major breakthroughs in understanding how the human brain functions, which will provide possible solutions to curing brain-related diseases. These projects will also inspire neuromorphic computing to meet growing demand for artificial intelligence² to build a machine that mimics the capability of humans towards various applications.

To obtain a better understanding of the brain, investigations of biological neural networks (BNNs) have been widely carried out to study the biological, structural and functional features of the brain. One category of this research involves the use of advanced electrophysiology and imaging techniques to map and study the activities of BNNs, which include: microelectrodes³, electroencephalograms (EEGs)⁴, magnetic resonance imaging⁵, computed tomography⁶, electron beam microscopy⁶ and super-resolution fluorescence microscopy⁷. This category allows an understanding of the operation of the brain via a bottom–up approach using the large amount of data obtained through BNN studies. Another category of BNN research focuses on the building of artificial neural networks (ANNs) that could emulate the biological, structural and functional features of BNNs. This category uses simplified and controllable models to test new theories of brain functions derived from the data of BNNs and to test new drugs on the brain-related diseases. ANNs can also provide brain-like computing platforms for artificial intelligence with higher efficiencies.

Building ANNs includes research in the fields of software simulations based on conventional von Neumann computers⁸,⁹ and hardware simulations, such as the implementation of ANNs indirectly based on electronics¹⁰,¹¹ and photonics¹²–¹⁵ and the direct growth of ANNs with biological neuron cells¹⁶. Although versatile supercomputers based on conventional von Neumann computers are now available to conduct millions of operations, the software simulation of ANNs with von Neumann computers in 100% real time and at the scale of the whole human brain has not yet been achieved and would consume at least 500 MW of energy¹⁷, not to mention its huge size. These drawbacks are mainly owing to the serial nature of von Neumann computers, which is fundamentally different from how the BNNs work. The
development of ANNs based on hardware is an important step for the practical realisation of ANNs.

Nanophotonics—the study of the behaviour of light and the light–matter interaction at the nanometre scale—is of great importance to demonstrate hardware-based ANNs. In fact, the interplay between nanophotonics and ANNs has already generated new fields. The application of software-based ANNs in nanophotonics has enabled new realms in automatic optical sensing, automatic optical microscopy imaging and the inverse design of photonic devices. Nanophotonics is a highly promising tool for studying BNNs with optical imaging and optogenetics. The development of optical imaging—especially the invention of far-field super-resolution optical microscopy (awarded the 2014 Nobel Prize in Chemistry), has initiated research into imaging neural networks at the nanoscale. Based on the same principle as far-field super-resolution imaging, a two-beam optical nanolithography technique, termed super-resolution photoinduction-inhibited nanolithography (SPIN), has been developed to fabricate features as small as 9 nm, which reaches the feature size of BNNs. Such a technique has the potential to develop the three-dimensional (3D) direct nanoprinting of ANNs with high complexity and capacity. In addition, the photon is an important information carrier for ANNs, exhibiting a broad bandwidth and low transmission scattering compared with the electron. The development of ANNs based on nanophotonic devices has opened a new avenue to achieve orders-of-magnitude improvements in computational speed and energy consumption over existing solutions based on electronics (Fig. 1).

In the present review, we discuss the recent advancement in building ANNs based on biological neurons and the development of nanophotonic-enabled indirect ANNs based on electronics and photonics, including the laser writing of electronic devices for ANNs and the development of nanophotonic devices for ANNs. After that, we briefly summarise the recent advancement in building ANNs based on biological neurons with controlled topology. Finally, we summarise the nanophotonic techniques used for the imaging and signal detection of ANNs based on biological neurons.

Indirect ANNs enabled by nanophotonics

The construction of intelligent machines that mimic BNNs has been pursued since soon after the invention of the modern computer. Most of the research into ANNs is based on software simulation using von Neumann computers. The concept of mimicking BNNs with electronic or photonic hardware, which is also called neuromorphic computing, was introduced in late 1980s. Compared with the ANNs directly based on biological neural cells, these ANNs rely on electronic or photonic systems containing artificial neurons to indirectly mimic the neurobiological architectures presented in BNNs.

Working principles of indirect ANNs

The biological counterparts of artificial neurons are biological neurons, which represent the building blocks of BNNs. The BNNs consist of billions of neurons of different types and sizes. Figure 2a shows a schematic of a simplified biological neuron with the four basic fundamental units—dendrites, axons, somas and synapses.

ANNs follow a simplified model inspired by BNNs. The unit of this model, a simple formalised artificial neuron introduced by McCulloch and Pitts, acts as a computational element in the ANN. The execution of a task involves the parallel activation of a large number of artificial neurons. As shown in Fig. 2b, one or more inputs (x) from the other neurons were sent into an artificial neuron and the inputs are summed up to produce outputs (y) to the other neurons on axons. Separated by weighted (w), a non-linear function known as an activation function or transfer function (f); the activation function can be a step function, Sigmoid function, etc. The mathematical form of the artificial neuron is as follows:

\[ y_j = f \left( \sum_{i=0}^{n} w_i x_i \right) \]  

The functions of the key components (axon, dendrite, soma and synapse) in ANNs are summarised in Table 1. The analogy between artificial neurons and biological neurons holds that the interconnections of the signals stand for the axons and dendrites, the summation and threshold approximate the activation in the soma, and the connection weights and memory represent the synapses.

Artificial neurons can be organised in any topological architecture to form ANNs. A common architecture, the
feed forward ANN, is shown in Fig. 2c. In this network, the information moves only in the forward direction. A series of artificial neurons are integrated to form the layer of input. Within the ANNs, the input signal is passed through one or more hidden layers. At the end of this structure, the output layer provides the results. Different from the current von Neumann computer that performs tasks by the pre-design of the programme, the ANN can learn to run a task through a sequence of training with examples. For example, one of the typical learning processes is the supervised learning method by updating weights with the backpropagation errors between known target values and the output values.

The most common unsupervised learning method is the spike timing-dependent plasticity (STDP) algorithm, inspired from the spiking nature of BNNs. Recent neurological research has shown that biological neurons encode information in the timing of single spikes. Functionally, STDP constitutes a mechanism for implementing a Hebbian learning rule, allowing for the learning process without supervision.

Laser-written electronic memristors for ANNs based on electronics

The ANNs based on electronics were first demonstrated with very large-scale integration (VLSI) systems. The electronic wires on the chips can be used as electronic axons and dendrites. The combination of several transistors in VLSI are used to approximate electronic synapses and somas. ANNs with a large amount of artificial neurons based on VLSI have been developed. However, the integrated level of such a kind of ANN is still low owing to the complexity of the architecture. Memristors, the fourth basic element in electronic circuits predicted in 1971 and demonstrated in 2008, provide a new path to demonstrate ANNs with a high integration of electronic devices.

Memristors are passive two-terminal circuit elements, usually with a metal/insulator/metal structure. Mimicking a biological synapse, the resistance of a memristor can be adjusted by tuning input amplitude of the charge or flux. It has been demonstrated that the synaptic functions can be implemented experimentally.

### Table 1 Summary of the key components of artificial neural networks

| Components       | Functions                      | Photonic devices                                      |
|------------------|--------------------------------|-------------------------------------------------------|
| Axon and dendrite| Interconnection                | Free space, waveguides                                |
| Soma             | Summation and thresholding     | Photodetectors, electric-optic modulators, light sources (LED and laser), optical amplifiers, saturable absorbers, optical bistable devices |
| Synapse          | Weighting                      | Hologram, MZI, micro-ring resonators                  |
|                  | Memory                         | Electronic memories, reversible optical memories      |
with nanoscale silicon-based memristors. The learning function has been demonstrated in a circuit composed of memristor synapses and complementary metal-oxide semiconductor neurons. An artificial soma function has been demonstrated with a neuristor built using two nanoscale Mott memristors. An ANN unsupervised learning for pattern classification based on memristors has recently been demonstrated. Nanophotonics offers a new solution to fabricate memristors by the laser reduction of graphene oxide. Graphene has received significant attention owing to its superior electronic and optical properties. Direct laser writing (DLW) with laser reduction can realise the simultaneous direct growth and patterning of rGO. Various rGO electronic devices such as capacitors, gas sensors and multifunctional devices have been demonstrated.

A memristor based on rGO has been experimentally demonstrated. DLW is used to realise rGO fabrication for use as the bottom electrode of the memristor on flexible substrates. An insulating layer of 10 nm HfOx is blanket deposited by thermal evaporation. The top electrode is made by curable silver paste or using thermal evaporation. The fabricated Ag/HfOx/rGO structure exhibits stable switching up to 100 cycles. Metal-free memristors have been fabricated through the DLW process with rGO/g-C3N4-NSs/rGO thin films.

Introduction of ANNs based on photonics

Photons are an ideal information carrier because of their specific properties, such as inherently massive parallelism, fast propagation speed and no side effects of mutual interference. Optical signals can be multiplexed in time, space, polarisation, angular momentum and wavelength domains, and optical technologies may overcome the problems inherent to electronics. Thus, the research and development into optical interconnect technology has already led to the replacement of copper connections in computer chips and data centres by optical waveguides or fibres, which can potentially improve the performance of ANNs based on software simulation and electronics. An ANN based on reservoir computing has been demonstrated with temporal multiplexing in a single optical fibre. ANNs based on nanophotonics offer a promising alternative approach with a faster response and lower power consumption.

ANNs based on nanophotonics can be achieved by providing key optoelectronic or optical components for the functions of ANNs (Table 1). Nanophotonic technology has several advantages in making interconnections with free space and waveguides, specifically with regard to broad bandwidth, low-loss and low-crosstalk. Photonic somas can be achieved either by optical–electrical and electrical–optical signal transferring with photodetectors, electro-optic modulators and light sources (LED and laser) or by an all-optical process with lasers, optical amplifiers, saturable absorbers and optical bistable devices. The weighting function can be achieved by optical switches such as holograms, Mach–Zehnder interferometers (MZIs) and MRRs. The memory function to record the weights in a synapse can be achieved by electronic memories or non-volatile optical memories.

ANNs based on holography in free space

In ANNs based on holography, the full parallelism of light in free space can be used and additional functionalities become available. These strengths have been recognised for many years, and early implementations utilised reconfigurable holograms for forming interconnections between optoelectronic artificial neurons. The architect of the ANNs based on holography with optoelectronic artificial neurons is shown in Fig. 4a. The light sources and photodetectors combine as optoelectronic neurons. The input signal is generated by the modulation of the light source. The holographic interconnection is achieved by optical diffraction with...
holographic gratings. Holography gratings can be produced by using spatial light modulators (SLMs) and the photorefractive effect. The detectors integrate the optical signal and transfer it into an electronic signal. The threshold function is achieved by processing the electronic signal with electronic circuits. Although free-space
optics provide the media for the highly integrated large number of free-space interconnections and data communication at the speed of light, the hologram provides an approach for storing and implementing the weight matrix data between fully connected neural layers.

Recently, a all-optical deep learning framework in the terahertz region (0.4 THz) has been demonstrated as an all-optical deep learning network, in which the ANN consists of multiple layers of a hologram fabricated by 3D additive printing. The ANN is created by using several layers of a diffractive surface, where each point on a given layer represents an artificial neuron (Fig. 4b). The learning process of such an ANN is based on a computer simulation to design the phase modulation of each layer (Fig. 4c), which can be fabricated with a 3D printer and based on plastic materials. The ANN cannot be tuned after the fabrication process. Compared with previous ANNs based on holography composed of optoelectronic neurons, this kind of ANN is based on all passive elements.

The ANNs based on holography can be easily scaled up using various fabrication techniques including electron beam lithography and photolithography. ANNs based on holography provide the potential for a large number of artificial synapses and neurons owing to the high parallelism in free space. If we assume that the artificial neurons in two adjacent layers of an ANN are all arbitrary interconnected. The relationship between the density of artificial neurons \(D_n\) and the density of artificial synapse \(D_s\) is:

\[
D_n = \frac{4}{\pi} \cdot \sqrt{D_s}
\]

The maximum number of artificial synapse via holography is equal to the maximum number of gratings that can be supported by the hologram. From the sampling theorem, the number of sinusoids or gratings that can be recorded in the medium is equal to the number of samples. Considering the pixel size of a hologram is \(\delta\) in all dimension \((\delta > \lambda)\), the maximum number of independent artificial synapse is \(V/\delta^3\) in a volumetric hologram or \(A/\delta^2\) in a thin hologram, where \(V\) and \(A\) are the volume and area of the hologram, respectively. As shown in Fig. 4d, for a commercial SLM with a pixel size of 4 \(\mu m\), the areal density of synapses \((D_{na})\) is \(6.25 \times 10^4/mm^2\). A recent advancement in holography induced by laser has demonstrated a hologram with a pixel size of 0.55 \(\mu m\), which corresponds to a synapse density of \(3.31 \times 10^6/mm^2\). As this hologram is induced by the tight focusing of laser light, it has potential to achieve a volume synapse density \((D_{va})\) of \(6.01 \times 10^7/mm^3\) if a volume hologram can be induced by laser in three dimensions. The corresponding maximum areal neuron density \((D_{na})\) and volume neuron density \((D_{va})\) can be calculated by Eq. (2), as shown in Fig. 4d. A recent advancement in the SPIN technique, with a feature size of 9 nm and optical signal multiplexing, can further improve the density of synapses and neurons.

Apart from plastic materials and SLMs, ANNs based on holography can be achieved with different materials (Table 1). The 3D two-photon polymerisation of polymers and chalcogenide glass can be used for ANNs based on holography. Holography imaging has been demonstrated by DLW with the photoreduction of graphene oxide. A hologram with a thickness of 60 nm using a topological insulator material has been demonstrated by DLW. ANNs based on the materials above are write-once only owing to the non-reversible photoinduced effect, requiring learning with pre-design structures in computers.

Reversible holograms have been demonstrated with photorefractive polymers and photorefractive polymers synthesised with nanoparticles. A hologram based on an ultrathin layer of phase change material Ge\(_2\)Sb\(_2\)Te\(_5\) (GST) is also demonstrated. These holograms can be erased and written by the DLW technique owing to the thermally reconfigurability of GST. These reversible holography effects open a new avenue for ANNs with a closed-loop learning capability, which allows direct updates of the weighting in ANNs after receiving feedback from the outputs.

**ANNs based on photonic circuits on a chip**

Integrated photonic circuits on a chip are an ideal platform for ANNs with a high compactness and high stability. As the photonic axons and dendrites, the low-loss waveguides transmit the optical signal. The development of integrated lasers, photodetectors and optical non-linear devices can serve as photonic somas. The weights can be achieved in the optical domain by using reconfigurable optical switches or splitters, which utilise the temperature dependence of the refractive index (the thermo-optics effect) to realise reconfigurable functionality by heating the devices in most of current devices.

Interest in integrated lasers with neuron-like-spiking behaviour has flourished over the past several years. Biological neurons use spikes to send information. Recently, neurological research has discovered that information in biological neurons are encoded in the temporal domain of single spikes. Lasers operating in the excitable regime are perfectly dynamic candidates but are ~ 8 orders-of-magnitude faster. However, most experimental work has focused on isolated neurons or single chain of neurons (Fig. 5a). Isolated artificial synapses on photonic circuits have been demonstrated by integrating a phase change memory (PCM) on an optical waveguide. As shown in Fig. 5b, the synapse is based on a tapered waveguide with PCM regions on top. The synaptic weight can be controlled by changing the number of optical pulses sent through the waveguide.

In parallel with the work on isolated artificial neurons and synapses, recent research has also investigated the
capability of photonic circuits to mimic neural networks, focusing on the networks with connectivity to multiple neurons and layers. An ANN scheme based on silicon waveguide interconnections has been demonstrated by using a reconfigurable MZI to control the weights between the interconnections of neurons (Fig. 5c), which are laser and photodetectors. Recently, a “broadcast-and-weight” scheme has been demonstrated to use wavelength-division multiplexing (WDM) to support a large number of reconfigurable interconnections enabled by micro-ring resonators (MRRs) on a silicon photonic chip. A 294-fold acceleration against a conventional von Neumann computer is predicted.

The size of the reconfigurable MRRs or MZIs is from 625 to 20,000 μm², resulting in a synapse density of 50–1600/ mm². The neuron density is at the same level and is based on the size of the integrated lasers and photodetectors on the chip. As mentioned in the previous paragraphs, current ANNs utilise the thermo-optics effect in reconfigurable MZIs or MRRs to realise artificial synapse functionality by heating the devices, resulting in an additional power consumption of 10 mW per synapse to maintain the reconfigurable MRRs or MZIs. However, the weights can be set by integrated on-chip non-volatile memory, such as a PCM, which requires no power to maintain.

**Future perspectives**

The ANNs based on photonics have a significantly higher rate and lower computation energy compared to those of their electronic counter parts. ANNs based on holography can achieve an ultrahigh density of synapses owing to the high parallel processing in free space. Nonetheless, current ANNs based on photonic circuits cannot compete with the density of electronics. However, novel implementations of ANNs based on nanoscale optical data storage, photonic crystal nanocavity and plasmonic nanocavity technologies may become important to the future scaling down of the artificial neuron and synapses on a chip to the diffraction limit scale (~ 1 μm) or even beyond the diffraction limit. Three-dimensional photonic integration enabled by DLW could enable ANNs with larger number of artificial neurons and synapses by adding another spatial degree of freedom. Furthermore, the feeding of input signals through temporal multiplexing would be able to realise ANNs with larger number of artificial neurons and synapses effectively.

Apart from the demonstration of single lasers or laser chains, the coding of ANNs based on photonics has been demonstrated using analogue optical signals. ANNs based on photonics can be coded with the spiking method to further improve the signal robustness and enable unsupervised learning. Current ANNs based on photonics are derived from simplified artificial neuron models. The performance of ANNs can be further improved with a better understanding of the working mechanism of BNNs.

**Direct ANNs enabled by nanophotonics**

One of the fundamental interests in building ANNs lies in the possibility of unravelling the myth in the
interactions between neurons. Building well-defined ANNs based on biological neurons with controlled topology (micro-platforms)\textsuperscript{66} can be used for studying neural activities such as axonal path finding and synaptogenesis, drug screening and targeting, and recreation computational units based on living cells\textsuperscript{67}. Furthermore, transplantable ANNs\textsuperscript{68} consisting of biological neural tissue might offer the capability to treat several diseases that interrupt the connectome in neural systems, including Parkinson’s disease, traumatic brain injury, stroke and brain tumour excision\textsuperscript{69}. A crucial unit for the successful building of ANNs based on intricate 3D bio-tissue is the manufacturing and adoption of well-designed 3D scaffolds. Successful scaffolds in tissue engineering applications rely critically on the physical and chemical characteristics of microstructures. Nanophotonics is the key enabling technology for the building of 3D scaffolds. In this chapter, we summarise different approaches and materials that are available to build ANNs enabled by nanophotonics.

**Fabrication of direct ANNs in two dimensions**

The mechanical and geometrical features of the surrounding structural matrix can have an impact on the structure and functions of the neurons\textsuperscript{20}. To create well-defined direct ANNs with a desired topology, micro-platforms that emulate the structural features of BNNs should be created on a biologically compatible substrate, as shown in Fig. 6a–c. Different lithographic techniques have been utilised to create 2D microstructures on biologically compatible materials.

Soft lithography can be used to build small hutch to trap neurons on the surface of electrodes\textsuperscript{71}. Similar structures such as troughs and wells can be etched into a silicon wafer. A triangular network with next-nearest-neighbour connections can be manufactured with the same method\textsuperscript{72}. The biocompatible materials for ANN fabrication are polydimethylsiloxane (PDMS) parylene\textsuperscript{73} and hydrogels\textsuperscript{74}. With structures like through-holes, neuron cells can be immobilised and connected and using micro tunnels, which can be encoded into a layer of

**Fig. 6 Recent developments in building artificial neuronal networks based on biological neurons with controlled topologies at different scales using different techniques**

A–c Steps in the creation of 2D ANNs using 2D fabrication technologies\textsuperscript{72} (photolithography, electron beam lithography, and micro-imprinting). d–f Millimetre-scale 3D ANNs created by 3D additive printing\textsuperscript{79}. g–i The creation of micrometre scale 3D ANNs using 3D direct laser writing\textsuperscript{87}.
PDMS or hydrogel for neurite outgrowth nanodiamond (ND) layering\textsuperscript{75} has also been demonstrated to be an excellent growth substrate for functional neuronal networks with the aid of photolithography methods.

After the fabrication of a scaffold, the ANN with desired topology can be achieved by cell placement, which is usually the combination of cell delivery and localisation. Cell placement is usually feasible using glass micropipettes\textsuperscript{76}, microfluidics, laser guidance\textsuperscript{77} and micro-manipulation systems\textsuperscript{78} for pick-and-place positioning.

**Fabrication of direct ANNs in two dimensions**

The current methods for building direct ANNs based on biological neurons are largely limited to two-dimensional (2D) substrates. Such 2D neural cultures cannot achieve the 3D spatial extensions of axons and dendrites. However, BNNs in the human brain possess extraordinary connectivity and complexity from the millimetre-scale down to a scale of several nanometres in 3D, which include the microscale of single neurons and synapses (nm–\(\mu\)m, synapse-by-synapse), the mesoscale of neuronal populations and their interconnecting circuitry (region-by-region), and the macroscale of anatomically distinct brain regions and pathways. Consequently, the development of accurate ANNs that mimic the brain remains a significant obstacle to our understanding of the functioning of the brain at different levels.

Recent developments in using 3D additive printing technique demonstrated it great capability potential in building large-scale 3D ANNs (Fig. 6d–f). Hydrogel-made 3D brain-like structures consisting of primary neurons have been created with a peptide-modified biopolymer\textsuperscript{79}. This technique has been used for the creation of intricate functionalized 3D brain-like modular formed from cortical tissue, maintained alive for months in vitro. 3D architectures that compartmentalising biological tissues have also been achieved by 3D additive printing using silk-collagen protein scaffolds\textsuperscript{80} and hydrogels\textsuperscript{81}, which have then been used to build 3D brain-like tissue seeded with primary cortical neurons.

One clear disadvantage of these 3D ANNs is that the connections between neurons are randomly patterned. The fabrication resolution of traditional 3D additive printing limits its capability of creating 3D ANNs with micrometre- or nanometre scale resolution.

**Towards building direct ANNs based on biological neurons with nanoscale resolution in three dimensions**

One potential solution towards the building of 3D and nanometre scale ANNs lies in the recent development of 3D DLW. 3D DLW, such as single-beam two-photon DLW and two-beam SPIN\textsuperscript{23,82}, has been widely studied and utilised to produce 3D nanophotonic structures\textsuperscript{82}, holograms\textsuperscript{46}, microfluidics\textsuperscript{83}, biomedical implants\textsuperscript{84,85}, 3D scaffolds for cell cultures and tissue engineering\textsuperscript{86,87} and biomimetic neuron structures\textsuperscript{88–91}.

Owing to an intrinsic ability to produce 3D structures with a wide range of photo-sensitive materials, single-beam two-photon DLW has been used to fabricate scaffolds for ANNs with biological cells to study the growth of neurons and guidance\textsuperscript{92}, as shown in Fig. 6g–i. Low-profile barrier structures have been successfully fabricated using bovine serum albumin (BSA) and laminin to guide the interactions of brain cortical neurons and neuroblastoma–glioma hybrid cells (NG108-15) in neuron cell culture\textsuperscript{93,94}. Using hyaluronic acid hydrogels, guidance pathways of biotinylated BSA functionalized with IKVAV peptides for rat dorsal root ganglion cells and rat hippocampal neural progenitor cells have also been created\textsuperscript{95}. 2PP have also been used to create scaffold using synthetic biodegradable polymers. 3D structures have been fabricated using polylactide-based photo-polymer with the shape of linear guidewires, with which the directed growth behaviour of NG108-15 and PC12 neuroblastoma cells have been studied\textsuperscript{86}.

**Future perspectives**

Intense ongoing research activity is focused on the use of two-photon DLW in the fabrication of 3D scaffolds for neuron growth and neuron tissue regeneration. Two-photon DLW is the only technique able to provide the 3D fabrication capability and high spatial resolution that are necessary to create a scaffold mimicking the structural features of BNNs. This technique may offer the means to produce scaffolds that are bioactive at micrometre scale and even at nanometre scale.

Modification of the physical and chemical characteristics of 3D nano/microstructures with submicron resolution will act as a crucial key in the study of BNNs. Because of the unique electrical and optical properties assembled on these platforms, they also hold the potential to augment brain functions. Creation of 3D nano/microstructures-based biosensors is faced with many challenges, but the aim is to achieve single electrical connections in brain neuron circuits and neural networks of interest. Eventually, nano/microstructures manifest the interactive platform between nanotechnology and neuroscience, making them promising medium in neuron technology for diagnosing and treatment of brain diseases in neurology.

**Signal detection of direct ANNs enabled by nanophotonics**

Unlike indirect ANNs based on electronics and photonics with input and output connections for signal detections, the spiking of biological neurons are the signals of direct ANNs. The detection of such weak electric signals at the nanoscale is of critical importance to studying neural activities. Novel devices that are ultrasensitive to
an mw field at the zero-field splitting frequency, $D = 2.87$ GHz, there is a redistribution of the electrons from the most populated sublevel, $m_s = 0$, to the less-populated sublevel, $m_s = \pm 1$.

Under an excitation beam at wavelength 532 nm, an NV$^-$ centre is polarised in its ground state $m_s = 0$. The application of an mw field at the zero-field splitting frequency causes an increase in the population of the $m_s = \pm 1$ spin sublevels. This leads to an overall decrease in the fluorescence because of the non-radiative decay via the intermediate metastable state. The observed photoluminescence decrease is called the optically detected magnetic resonance (ODMR) signal. It is also possible to apply a magnetic field along the $z$ axis that degenerates the energy sublevels $m_s = \pm 1$.

In case of the weak-magnetic field regime, the magnetic field is calculated with the following\textsuperscript{104}:

$$B_{NV} = \frac{\omega_2 - \omega_1}{2g_s\mu B}$$

(4)

Under this circumstance, the recorded ODMR signal shows two dips owing to the redistribution of the electrons in the sublevels $m_s = -1$ and $m_s = +1$. The separation between the recorded ODMR dips increases with the strength of the applied magnetic field.

In fact, APs are time-varying electrical fields, which lead to time-varying magnetic fields. The main principle behind sensing APs with NV$^-$ centres is the detection of time-varying magnetic fields generated by APs via ODMR signals. The measurement of the AP via magnetic field sensing confers important advantages: it is non-invasive, label-free and able to detect neuronal activity through intervening tissue and whole organisms.

In the first work of applying NV$^-$ centres to measure the axon transmembrane potential, the magnetic field generated by a single axon potential is modelled, magnetic field generated by a single axon potential is modelled and the magnetic field has been generated by a microwire on the surface of a diamond substrate, which simulates the AP of a morphologically reconstructed hippocampal CA1 pyramidal neuron. The detection system is composed of a commercial grade single crystal ultra-pure diamond substrate with a layer NV$^-$ defect centres with a standoff of 100 nm. The detection system is able to image planar neuron dynamics non-invasively with temporal resolution at millisecond and spatial resolution at micron spatial resolution (10 $\mu$m$^{-1}$) within wide-field view\textsuperscript{105}. NV$^-$ centres in bulk diamonds have been implemented to measure the APs of a giant axon in an invertebrate, Myxicola infundibulum\textsuperscript{97}.

The bipolar azimuthal magnetic field associated with an AP is depicted in Fig. 7a, in the inset is the energy level of the NV$^-$ centres. A light beam at a wavelength of 532 nm is applied to the sensing NV$^-$ centre layer through the weak electric fields, such as micro-electrode arrays and patch clamps, have been introduced to study neural activities. However, these devices have a low spatial resolution and are invasive, which might cause damage to the biological neurons. Nanophotonics, especially nano-sensing enabled by nitrogen-vacancy (NV$^-$) centres in bulk diamonds and NDs, has opened the possibility of the far-field detection of neural activities with nanoscale resolution.

Sensing of the action potential

The electrical potential associated with the passage of a spiking impulse along neurons is called action potential (AP). APs play a central role in the communication between individual neurons. The signals between neurons propagate along axons via APs. Different techniques can be implemented for the measurement of APs. Electrophysiology is a recording method with a patch-clamping configuration that remains the gold standard for the measurement of individual APs. The technique has an excellent temporal resolution and good signal-to-noise ratio, but the spatial resolution is limited to ~ 10 µm.

Nanophotonic techniques to measure APs offer many advantages, but they typically require high power, which can cause photodamage to neurons. In addition, voltage-sensitive fluorescent proteins must be genetically expressed, which may alter the neuronal functions. Owing to their optical and magnetic properties, NV$^-$ centres can be applied to study neural networks and the firing of neural cells to study brain activity\textsuperscript{97}.

NV$^-$ centres have attracted significant attention in the recent years owing to their outstanding optical and magnetic properties\textsuperscript{98,99}. The ground state of an NV$^-$ centre is a triplet state with six unpaired electron spins, each one with an associated magnetic moment; the spin sublevels are $m_s = \pm 1, 0$\textsuperscript{100}. The energy sublevels $m_s = \pm 1$ in the ground state can be perturbed with an external magnetic field. This paramagnetism of the ground state can be mathematically described by the spin Hamiltonian of the system\textsuperscript{100}:

$$H = D_{gs}S_z^2 + E_{gs}\left(S_x^2 - S_y^2\right) + g_{gs}\mu B \cdot \vec{S}$$

(3)

where $D_{gs}$ is the ground state zero-field splitting, $\vec{S}$ is the spin operator vector, $E_{gs}$ is the ground state strain-induced splitting coefficient, $g_{gs}$ the ground state g-factor and $\mu$ is the Bohr magneton\textsuperscript{101}. In brief, the first term corresponds to the energy gap between the two sublevels in the ground state, the second term represents the energy shift owing to the strain from the lattice, and the last term is the Zeeman effect\textsuperscript{102,103}.

When NDs are stimulated by a microwave (mw) field, it is possible to redistribute the population of ground state electrons in NV$^-$ centres. In particular, by switching on
Diamond
1A1
3A2
1E

AP propagation

Specimen

AP magnetic field

Diamond

Specimen

Bipolar electrodes

Perfusion bath

Suction electrode

532 nm

light

NV fluorescence collection

Magnet

Conventional imaging objective

NV layer

Worm A axon

Measured Φ(t)

-20

-40

-60

-80

-100

0

24

Time (ms)

6

Calculated magnetic field (a.u.)

Calculated B(t)

0

24

Time (ms)

6

Measured action potential (mV)

Worm A axon

Measured Φ(t)

-100

-80

-60

-40

-20

0

24

Time (ms)

6

Calculated magnetic field (nT)

Worm A axon

Measured B(t)

-2

-1

0

1

2

3

24

Time (ms)

6

Nuclei

β-III-tubulin

SSEA-1

FND

Merge

ECSC cells (FND treated)

Differentiated neuronal cells (FND treated)

β-III-tubulin

SSEA-1

FND

Merge

Fig. 7 (See legend on next page.)
diamond at a sufficiently shallow angle so that the light reflects off the top diamond surface and therefore does not irradiate the living sample (Fig. 7b)\(^97\). The magnetic field from the AP measured via electrophysiology can be mathematically expressed as \(^97\):

\[
\phi_{\text{meas}}(t) = B(t) = \int sd\phi_{\text{int}}dt
\]

The measured AP voltage and the calculated magnetic field are illustrated in Fig. 7c–e. The measurements with NV\(^-\) centres are consistent with the prediction\(^97\).

With NV\(^-\) centres, it is also possible to measure the AP propagation direction at a single point\(^97\). The results lay the groundwork for the real time, non-invasive 3D magnetic mapping of functional neuronal networks, ultimately with a circuit-scale (~ 1 cm) field-of-view.

**Labelling and tracking of neuronal differentiation with NDs**

An understanding of biological processes involves the tracking of several cells to study their differentiation and development. Owing to their photostability, their robustness against photobleaching and their biocompatibility, NDs are ideal candidates for long-term cell tracking\(^\text{106}\). A study has been reported on the tracking of neuron cells derived from a model of embryonal carcinoma stem cells with fluorescent NDs\(^\text{107}\). The NDs are detected with a confocal microscope (Fig. 7f). No effect on the morphological development of the cells is recorded nor do the NDs induce apoptosis during the neuronal differentiation (Fig. 7g). Therefore, the implementation of fluorescent NDs to track neural cell development could provide potential therapeutics for neural diseases\(^\text{107}\).

**Future perspectives**

Studies on the detection of APs by measuring the ODMR signal with NV\(^-\) centres in bulk diamonds have been reported\(^\text{97}\). There are several technical challenges with this technique. First, the magnetic field sensitivity needs to be improved to enable AP measurement from individual mammalian neurons\(^\text{97}\). Second, single-point detection needs to be developed for the 2D and 3D mapping of APs in neuronal networks. Third, the high-resolution imaging of AP magnetic fields is also required.

NDs can provide a new solution to the three challenges listed above. NDs have a wide range of applications in the life sciences in the life sciences compared with bulk diamonds\(^\text{108}\). Currently, NV\(^-\) centres in NDs have been applied to study the formation and patterning of neural networks and to label and track the neuron cells\(^\text{5,107}\). First, a higher density of NV\(^-\) centres can yield a sensitivity improvement. NDs provide a substantially higher density of NV\(^-\) centres than does bulk diamond. Some types of NDs contain ~ 1000 NV\(^-\) centres in one ND with a diameter of 5 nm. Second, NDs enable 3D magnetic imaging. The magnetic imaging based on diamond layers can only achieve 2D imaging because the neuron networks can only be fixed on the surface of the diamond layers. However, NDs can be labelled everywhere inside the neuron networks, which makes it possible to obtain 3D magnetic images. Third, a higher spatial resolution can be achieved with NDs. Smaller NDs in biomedical imaging are highly required, and the blinking phenomenon is more prevalent when reducing the size of a ND is reduced\(^\text{109–111}\). Super-resolution imaging based on blinding fluorescence would allow the nanometric reconstruction of synapse connections by labelling the neuron cells with 5 nm oxide NDs\(^\text{109,111}\).

The employment of the ODMR signal to interpret APs via magnetic field measurements could lay new groundwork for the development of nanoscale biomarkers for both the super-resolution optical imaging and magnetic sensing of ANNs.

**Conclusions**

In the present review, we have surveyed the recent advancements in nanophotonic technology for indirect and direct ANNs. Holography and integrated photonic circuits have shown potential for achieving indirect ANNs. To date, the integration density of functional devices in ANNs based on photonic circuits remains limited. Three-dimensional photonic integration enabled by DLW could be a solution to the development of indirect ANNs with a high bandwidth and low-power consumption. DLW can be used for the direct fabrication of ANNs in three dimensions. In addition, photonic sensing techniques can be used to study the neural activities in direct ANNs based on biological neurons.
Further developments in nanofabrication and super-resolution optical sensing techniques are essential to achieve fabrication and detection capabilities at the scale of the ANN nanofeatures.

Acknowledgements
We acknowledge the support from Australian Research Council (ARC) (Discovery Project 170101775).

Author contributions
MG proposed the framework of this review. All the authors contribute to the writing of manuscript.

Conflict of interest
The authors declare that they have no conflict of interest.

Received: 12 November 2018 Revised: 7 March 2019 Accepted: 26 March 2019
Published online: 08 May 2019

References
1. Reardon, S. Worldwide brain-mapping project sparks excitement—and concern. Nature 537, 597 (2016).
2. Scholkopf, B. Artificial intelligence: Learning to see and act. Nature 518, 486–487 (2015).
3. Rutten, W. L. C. Selective electrical interfaces with the nervous system. Annu. Rev. Biomed. Eng. 4, 407–452 (2002).
4. Del Pozo-Banos, M. et al. Electroencephalogram subject identification: a review. Expert Syst. Appl. 41, 6537–6554 (2014).
5. Cox, R. W. AFNI: Software for analysis and visualization of functional magnetic resonance images. Comput. Biomed. Res. 29, 162–173 (1996).
6. Makela, T. et al. A review of cardiac image registration methods. IEEE Trans. Med. Imaging 21, 1011–1021 (2002).
7. Berring, S. et al. Nanoscopics in a living mouse brain. Science 335, 551 (2012).
8. Markram, H. The blue brain project. Nat. Rev. Neurosci. 7, 153–160 (2006).
9. Hines, M. L. & Carnevale, N. T. The NEURON simulation environment. Neural Comput. 9, 1179–1209 (1997).
10. Merolla, P. A. et al. A million spiking-neuron integrated circuit with a scalable communication network and interface. Science 345, 687–693 (2014).
11. Schemmel, J. et al. A wafer-scale neuromorphic hardware system for large-scale neural modeling. Proceedings of 2010 IEEE International Symposium on Circuits and Systems. Paris, France: IEEE, pp.1947–1950 (2010).
12. Psaltis, D. et al. Holography in artificial neural networks. Nature 343, 325–330 (1990).
13. Tait, A. N. et al. Broadcast and weight: an integrated network for scalable photonic spike processing. J. Light Technol. 32, 4029–4041 (2014).
14. Tait, A. N. et al. Neuromorphic photonic networks using silicon photonic weight banks. Sci. Rep. 7, 7430 (2017).
15. Shen, Y. C. et al. Deep learning with coherent nanophotonic circuits. Nat. Photonics 11, 441–446 (2017).
16. Mahoney, M. L. & Ananth, K. S. Three-dimensional growth and function of neural tissue in degradable polyethylene glycol hydrogels. Biomaterials 27, 2265–2274 (2006).
17. Markram, H. et al. Reconstruction and simulation of neocortical microcircuitry. Cell 163, 456–492 (2015).
18. Zhang, K. W. et al. A fiber optic sensor for the measurement of surface roughness and displacement using artificial neural networks. IEEE Trans. Instrum. Meas. 46, 899–902 (1997).
19. Khan, J. et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat. Med. 7, 673–679 (2001).
20. Pearlfrey, J. et al. Nanophotonic particle simulation and inverse design using artificial neural networks. Sci. Adv. 4, eaar2066 (2018).
21. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission stimulated-emission-depletion fluorescence microscopy. Opt. Lett. 19, 780–782 (1994).
22. Betzig, E. et al. Imaging intracellular fluorescent proteins at nanometer resolution. Science 313, 1642–1645 (2006).
23. Gan, Z. S. et al. Three-dimensional deep sub-diffraction optical beam lithography with 9 nm feature size. Nat. Commun. 4, 2061 (2013).
24. Mead, C. Neuroorphic electronic systems. Proc. IEEE 78, 1629–1636 (1990).
25. Uhlig, R. E. Introduction to artificial neural networks. Proceedings of ICON ’95 – 21st Annual Conference on IEEE Industrial Electronics. Orlando, FL, USA: IEEE, pp.33–37 (1995).
26. McCulloch, W. S. & Pitts, W. A logical calculus of ideas immanent in nervous activity. Bull. Math. Biophys. 5, 115–133 (1943).
27. Hinton, G. Mental simulation. Nature 347, 627–628 (1990).
28. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. Nature 323, 533–536 (1986).
29. Freake, R. C. & Dan, Y. Spike-time-dependent synaptic modification induced by natural spike trains. Nature 416, 433–438 (2002).
30. Held, D. O. The organization of behavior in Neurocomputing: Foundations of Research (eds. Anderson, J. A. & Rosenfeld, E.) 43–54 (Cambridge, MA, USA: MIT Press, 1988).
31. Chua, L. Memristor—the missing circuit element. IEEE Trans. Circuit Theory 18, 507–519 (1971).
32. Stukov, D. B. et al. The missing memristor found. Nature 453, 80–83 (2008).
33. Williams, R. S. How we found the missing memristor. IEEE Spectr. 45, 28–35 (2008).
34. Jo, S. H. et al. Nanoscale memristor device as synapse in neuromorphic systems. Nano Lett. 10, 1297–1301 (2010).
35. Piazza, M. et al. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. Nature 521, 61–64 (2015).
36. Pickett, M. D., Medeiros-Ribeiro, G. & Williams, R. S. A scalable neuristor built with Mott memristors. Nat. Mater. 12, 114–117 (2013).
37. Wang, Z. R. et al. Fully memristive neural networks for pattern classification with unsupervised learning. Nat. Electron. 1, 137–145 (2018).
38. Geirn, A. K. & Novoselov, K. S. The rise of graphene. Nat. Mater. 6, 183–191 (2007).
39. El-Kady, M. F. & Kaner, R. B. Scalable fabrication of high-power graphene micro-supercapacitors for flexible and on-chip energy storage. Nat. Commun. 4, 1475 (2013).
40. Strong, V. et al. Patternning and electronic tuning of laser-scribed graphene for flexible all-carbon devices. ACS Nano 6, 1395–1403 (2012).
41. Tian, H. et al. Wafer-scale integration of graphene-based electronic, optoelectronic and electroacoustic devices. Sci. Rep. 4, 3598 (2014).
42. Tian, H. et al. Cost-effective, transfer-free, flexible resistive random access memory using laser-scribed reduced graphene oxide patterning technology. Nano Lett. 14, 3214–3219 (2014).
43. Zhao, F. et al. Functionalized graphene nanoribbon for metal-free, flexible and rewritable nonvolatile memory device via direct laser-writing. Sci. Rep. 4, 5882 (2014).
44. Deng, R. R. & Liu, X. G. Optical multiplexing: Tunable lifetime nanocry stal. Nat. Photonics 8, 10–12 (2014).
45. Zifjasta, P., Chon, J. W. M. & Gu, M. Five-dimensional optical recording mediated by surface plasmons in gold nanorods. Nature 459, 410–413 (2009).
46. Li, X. P. et al. Athermally photothesized graphene oxide for three-dimensional holographic images. Nat. Commun. 6, 6884 (2015).
47. Ren, H. R. et al. On-chip noninterference angular momentum multiplexing of broadband light. Science 352, 805–809 (2016).
48. Deng, R. R. et al. Temporal full-colour tuning through non-steady-state upconversion. Nat. Nanotechnol. 10, 237–242 (2015).
49. Asghari, M. & Krishnamoorthy, A. V. Silicon photonics. Energy-efficient communication. Nat. Photonics 5, 268–270 (2011).
50. Appeltant, L. et al. Information processing using a single dynamical node as complex system. Nat. Commun. 2, 486 (2011).
51. Mesaritakis, C. et al. Artificial neuron based on integrated semiconductor quantum dot mode-locked lasers. Sci. Rep. 6, 39317 (2016).
52. Rosenbluth, D. et al. A high performance photonic pulse processing device. Opt. Express 17, 22767–22772 (2009).
53. Li, S. H. & Cai, X. H. High contrast all optical bistable switching in coupled nonlinear photonic crystal microcavities. Appl. Phys. Lett. 96, 131114 (2010).
54. Rios, C. et al. Integrated all photonic non-volatile multi-level memory. Nat. Photonics 9, 725–732 (2015).
55. Kaikhah, K. & Loochan, F. Computer generated holograms for optical neural networks. Appl. Intell. 14, 145–160 (2001).
