



Drawing inspiration from biological dendrites to empower artificial neural networks

Spyridon Chavlis and Panayiota Poirazi

Abstract

This article highlights specific features of biological neurons and their dendritic trees, whose adoption may help advance artificial neural networks used in various machine learning applications. Advancements could take the form of increased computational capabilities and/or reduced power consumption. Proposed features include dendritic anatomy, dendritic nonlinearities, and compartmentalized plasticity rules, all of which shape learning and information processing in biological networks. We discuss the computational benefits provided by these features in biological neurons and suggest ways to adopt them in artificial neurons in order to exploit the respective benefits in machine learning.

Addresses

Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Heraklion, 70013, Greece

Corresponding author: Poirazi, Panayiota (poirazi@imbb.forth.gr)

Current Opinion in Neurobiology 2021, **70**:1–10

This review comes from a themed issue on **Computational Neuroscience**

Edited by **Julijana Gjorgjieva** and **Ila Fiete**

For a complete overview see the [Issue](#) and the [Editorial](#)

<https://doi.org/10.1016/j.conb.2021.04.007>

0959-4388/© 2021 Elsevier Ltd. All rights reserved.

Introduction

Artificial neural networks (ANN) implemented in deep learning (DL) architectures have been extremely successful in solving challenging machine learning (ML) problems such as image [1], speech [2], and face [3] recognition, playing online games [4], autonomous driving [5], and so forth. However, despite their huge successes, state-of-the-art DL architectures suffer from problems that seem rudimentary to real brains. For example, while becoming true experts in solving a specific task, they typically fail to transfer these skills to new problems without extensive retraining—a property known as “transfer learning”. Moreover, when learning new problems, they tend to forget previously learned ones—a problem termed “catastrophic forgetting.”

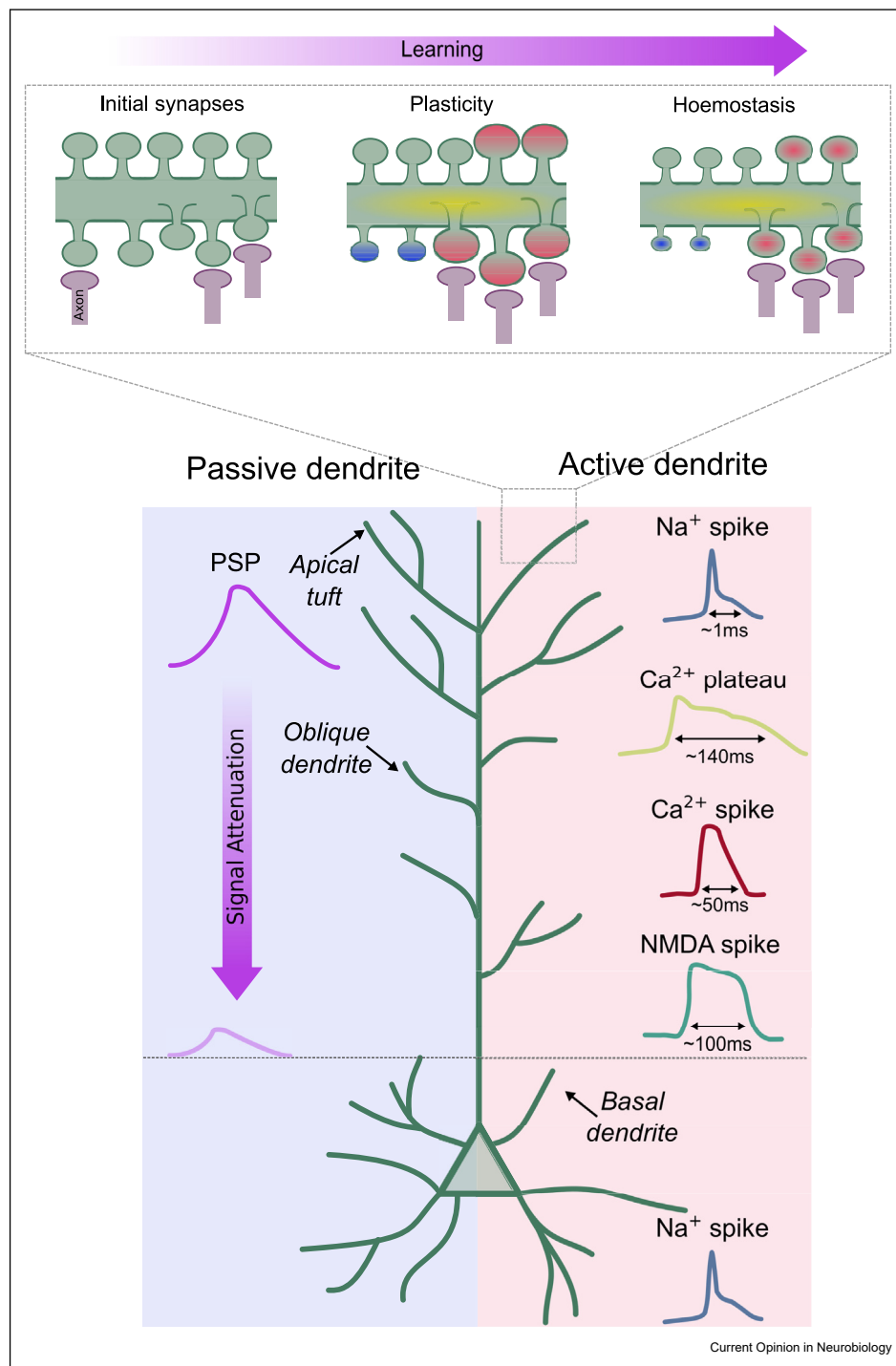
Finally, in order to achieve top performance, they typically require large training sets and thousands/millions of trainable parameters, necessitating the use of computing systems that leave a substantial carbon footprint. The brain, on the other hand, runs on extremely low energy (<20 W), learns from few examples, generalizes extremely well, and can learn in a continual manner without erasing useful prior knowledge. In this article, we argue that increasing the bioinspiration of ANNs, by adopting dendritic features may help address the abovementioned limitations. We discuss two types of benefits: advancing the computational power of ANNs and/or reducing their power consumption.

Why dendrites?

Dendrites are thin processes that extend from the cell bodies of neurons and serve as their primary receiving end: over 90% of the synapses in principal neurons are localized within their dendrites. Dendrites come in a great variety of shapes depending on the species, brain area, cell type, and even the layer in which they reside, indicating that they may be specialized for the respective computations needed [6]. Beyond their function as entry devices, morphologically complex dendrites affect the integration and propagation of incoming signals in complex ways. For example, passive dendrites act as filters due to the attenuation and temporal filtering of signals traveling to the cell body [7], thus allowing neurons to infer the distance of incoming signals from the soma and their kinetics. Moreover, active dendrites are home to a variety of voltage-gated conductances that support the generation of dendritic spikes (reviewed in Ref. [8], [Figure 1](#)). These dendritic features amount to different types of nonlinear integration and allow individual neurons to perform complex computations, ranging from logical operations to signal amplification or segregation, temporal sequence/coincidence detection, parallel nonlinear processing, and so forth [9–15].

The computational capabilities of neurons with nonlinear dendrites are exploited by a set of multiscale plasticity rules that induce synaptic, intrinsic, homeostatic, and wiring changes [16,17]. These rules can differ between neuronal (e.g., apical vs. basal trees) and/or dendritic (e.g., proximal vs. distal) compartments [18–22], enabling

Figure 1



Schematic representation of dendritic features in biological neurons. (Bottom) Schematic depiction of a typical pyramidal cell. If dendrites were passive (blue shaded area, left), the magnitude of an incoming signal would decrease as it propagates to the soma. However, with active dendrites (red-shaded area, right), a strong-enough incoming stimulus can generate various types of dendritic spikes ('dSpikes'). Typical shapes of 'dSpikes' are shown schematically on the right (blue: Na⁺ spike (duration: 1-2 ms), red: Ca²⁺ spike (duration: a few 10s of milliseconds), green: NMDA spike (duration: 50-150 ms), yellow: Ca²⁺ plateau potentials, which are mixtures of Ca²⁺ and NMDA spikes). (Top) Schematic representation of some plasticity mechanisms that operate in dendrites. (left) Initial connectivity, all synapses are of the same strength. (middle) Some of the synapses (the co-active ones) are strengthened, while others are depressed. The overall dendritic excitability increases (yellow shade). Axons that contact depressed spines rewire, and new connections are formed. (right) Homeostasis scales all synapses while maintaining their relative strengths.

neurons to alter dendritic integration in a targeted, nonuniform manner. For example, dendrites whose spikes do not propagate effectively to affect somatic output can turn into powerful influencers as a result of branch-strength-potential, a form of local intrinsic plasticity [23].

Importantly, the effects of anatomical, biophysical, and plasticity properties of dendrites in the living animal are often tangled. For example, considering only the anatomy of the dendritic tree, distant inputs should be severely attenuated before reaching the soma, causing weak somatic depolarization. However, this is not always the case. Voltage- and ligand-gated channels in the distal dendrites are activated upon the arrival of synchronous inputs, leading to the generation of dendritic spikes that reliably propagate to the soma, often causing somatic spikes (Figure 1). Because of these nonlinear interactions, biological neurons are considered much more powerful than linear ANN nodes (or point neurons) [12,24–27]. Instead, dendrites rather than neurons are proposed to serve as the fundamental functional unit of the brain [28]. Of note, the nonlinear computations performed by neurons with active dendrites can also be performed by networks of point neurons. However, solving them within individual neurons provides important savings: more energy would be required to fire APs across many neurons and to form appropriate axonal wiring/coordination of these neurons. Moreover, the space taken up by neuronal somata is much larger than the space occupied by dendrites; therefore, computing with dendrites is more resource-efficient for biological brains.

The potential benefits of the aforementioned dendritic characteristics have yet to be exploited in ANNs and their neuromorphic implementations. Respective neurons (nodes) in traditional ANNs are typically point neurons arranged in several layers, interconnected with symmetric (all-to-all), plastic (trainable) synapses, and furnished with nonlinear activation functions. Learning is typically achieved through weight updating rules, the most popular of which is the backpropagation-of-error algorithm. There are strengths in the simplicity of this approach (easier to understand and implement, fast convergence, high performance), but also weaknesses (e.g., catastrophic forgetting, high power consumption). To address the latter, researchers turn to more biologically inspired architectures [29–32] and learning rules [30,33,34], with promising results (e.g., on MNIST classification). We believe that an in-depth exploration of the anatomical, biophysical, and plasticity features of dendrites may help reduce the resources needed to achieve top performance and/or address problems like transfer and continual learning, where traditional ANNs fail.

Towards this goal, we review recent experimental and computational findings related to how dendritic features

can form a foundation for specific computations and suggest ways to integrate them in DL architectures. We focus on three fundamental features of dendrites: i) anatomy and compartmentalization, ii) active ionic mechanisms, and iii) multiscale plasticity rules. We spotlight recent discoveries that connect subcellular mechanisms and algorithmic-level functions for each one of these dendritic features.

Dendritic anatomy and compartmentalization

The pioneering work of Wilfrid Rall (1959) revealed that dendrites are not isopotential electrical devices but rather electrically distributed circuits whose conducting properties are described by the cable theory. As such, signals that travel along (passive) dendrites are logarithmically attenuated [35], and the level of this attenuation is highly dependent on the morphological properties of dendritic trees [7,36]. Specifically, the propagation loss is proposed to decrease with increasing diameter and decreasing length of dendritic compartments [37]. Given that dendritic trees are typically nonsymmetric, signal attenuation can vary extensively within the same neuron. Moreover, branch points are a major source of attenuation, a feature that promotes compartmentalized processing in neurons with elaborated trees [38]. These morphological effects have been attributed important functions: according to Rall's 1964 study [39], passive dendrites can detect the temporal sequence of incoming inputs by inducing different somatic responses depending on the activation order. This property could be used by neurons to infer the direction of motion (e.g., of a visual stimulus moving from right to left versus from left to right). Moreover, the compartmentalized, sublinear integration of inputs in passive dendrites was suggested to enable single neurons to compute linearly non-separable functions [40]. Finally, the dendritic structure has been shown to enhance sparsity in biological networks, a property associated with increased memory capacity [41], better discrimination [42], and faster learning [43]. For example, dendrites made granule cells less likely to fire in a dentate gyrus network model. Consequently, the network's ability to discriminate between similar input patterns was increased because patterns were encoded by less overlapping granule cell populations [42]. Moreover, sparsifying the connectivity of inputs, by distributing them across basal and apical dendrites, enabled simulated neurons to encode multiple input sequences simultaneously [44]. These examples highlight the power of dendritic anatomy in shaping signal processing in biological neurons, primarily by implementing sublinear integration and increasing network sparsity.

ANNs, on the other hand, typically use the McCulloch-Pitts model [45], whereby artificial neurons (nodes)

receive a weighted sum of their incoming synaptic inputs, largely ignoring dendritic influences. Yet, some of the approaches commonly used in the field to improve performance resemble the structural features of dendritic trees. For example, dropout layers used in deep architectures [46] lead to the sparsification of a network, as a randomly chosen portion of the connections are removed from training. Such sparsification is somewhat similar to that achieved by having multi-branched dendritic trees and achieves both resource savings (fewer trainable parameters) and better generalization performance. Modular networks [47,48], consisting of various ANN subnetworks, can also be seen as analogous to dendritic structures, with each module learning a part of the task. Finally, sparse ANNs, where a portion of the trainable parameters are set to zero (and respective multiplications can be omitted), have some similarities with the restricted connectivity imposed by dendritic structures [49].

Active ionic mechanisms and dendritic computations

In addition to their branching structure, many dendrites are equipped with a variety of voltage-gated and ligand-gated ionic channels. These so-called “active” mechanisms allow dendrites to transform incoming signals in nonlinear ways. For example, if two stimuli arrive within the same dendrite of pyramidal neurons, their integration results in a much larger response than if they land in separate dendrites [13,50,51]. This is because the integration of multiple nearby inputs causes a large enough depolarization that activates voltage-gated channels, which would otherwise remain closed, resulting in the generation of regenerative events (dendritic spikes). Dendritic spikes are much larger in amplitude than the linear summation of the incoming synaptic inputs (thus resulting in signal amplification) and can reliably propagate to the cell body, counteracting the attenuation caused by dendritic morphology. The ability to induce different responses depending on the dendritic location of inputs greatly increases the discriminatory power of neurons [24,52].

Dendritic spikes are driven by various ions and differ in their kinetics and localization. Dendritic sodium (Na^+) spikes are faster than calcium (Ca^{2+}) spikes and occur throughout the dendritic tree [10,11,53,54]. Ca^{2+} spikes are generated in the distal apical trunk and can spread to the apical dendritic tree [55,56]. NMDA spikes have even slower dynamics and usually reside within side/terminal dendritic branches [57]. Finally, Na^+ spikes generated at the soma can also travel back into the dendrites [58]. The occurrence of these dendritic spikes, their kinetics, and location dependence expand the computing power of individual neurons [59]. For example, both Na^+ and Ca^{2+} spikes have been attributed to a coincidence detector role. The former

were shown to emerge upon synchronous activation of synaptic inputs within an individual dendrite [10,50], thus enabling a neuron to detect synchronous inputs with millisecond precision. The latter were shown to emerge when distal and proximal dendritic regions of cortical neurons are activated within several milliseconds, leading to strong somatic bursting [60]. Bursting can signal the coupling (coincidence) of input streams from different regions, thus serving as an association mechanism operating at larger time scales. Finally, NMDA spike generation within individual branches was found to be sensitive to both the timing and location of inputs, thus endowing single dendrites with the ability to detect the activation sequence of inputs [14].

Beyond their spatio-temporal processing power, dendritic spikes were predicted by models [24,51] and verified by experiments [13,50] to enable dendrites to integrate inputs like sigmoidal activation functions. As a result, a pyramidal neuron with multiple dendrites (sigmoidal units) acts like a powerful 2-stage [24,25] or even a multistage ANN [61,62], with dendrites acting as computing nodes. Such dendritic nonlinearities were recently found in interneurons [63–65], suggesting that these cells may also act as multistage ANNs [27]. Notably, Gidon et al. [66] recently discovered a new, more powerful dendritic spike (dCaAP) in human layer 2/3 cortical neurons. Unlike other types of spikes, the input–output curve for dCaAPs is described by a nonmonotonic activation function. In biophysical models, this property enables a single human dendrite to solve the Exclusive-OR (XOR) problem, a logical operation though solvable only by multilayer ANNs. In sum, dendritic spikes endow dendrites with nonlinear signal processing capabilities, which enable the local implementation of various nonlinear computations and can be mathematically approximated with activation functions similar to those used in ANNs. Of note, the suppressing effects of passive dendrites (sublinear integration) can also be described with activation functions.

In the ML world, activation functions are a topic of intense investigation because of their enormous impact on the performance accuracy of ANNs [67]. Currently, the most popular activation function is the ReLU and its derivatives (e.g., PReLU, ELU, GELU), although more biological ones (e.g., sigmoid, tanh, delta, logarithmic) are also extensively used. The positive part of ReLU can be thought of as the frequency-current (F/I) curve of a neuron, while the negative part (e.g., leaky ReLU) is there for computational efficiency.

A number of recent ML studies have explored dendritic structure and activation functions in ANNs. For example [29], introduced three-compartment nodes comprised of a soma, an apical, and a basal dendrite, each furnished with its own activation function. Basal and apical

compartments received different types of information (sensory and feedback, respectively), allowing neurons in different layers to coordinate synaptic weight updates. As a result, the network learned to categorize images better than a single-layer network. While this work aimed at understanding how the compartmentalization of biological neurons may facilitate learning, it is one of the first studies where simplified dendrites were integrated as building blocks of the ANN graph. A similar approach using dendritic ANNs, that is, replacing each node with a 2-layered structure furnished with a dendritic and somatic layer, achieves higher model expressivity (complexity of functions that can be computed) compared to that of a single-layered neuron and better generalization than traditional ANNs, by sparsifying the network connections [68]. Jones and Kording [69] went a step further by incorporating multiple dendritic layers in an ANN. They showed that a single neuron modeled as a dendritic ANN with sparse connectivity could solve binary classification tasks (e.g., MNIST) and benefit from having the same input at different dendritic sites. Another approach utilizes multiple dendritic layers in each node of a deep ANN and reduces catastrophic forgetting in a continual learning task [70]. To explore the potential advantages of human dCaAPs, a recent study used their respective transfer function in various deep learning architectures and found a small but significant increase in classification performance for various tasks, especially when used in combination with sparse connectivity resembling that of dendritic trees [71]. It should be noted that none of the aforementioned studies has surpassed the current state-of-the-art in performance accuracy. However, it has been shown that dendritic ANNs can achieve better performance (lower error) when compared with traditional ANNs with the same number of trainable parameters [68].

While the true power of dendrites for ML is just beginning to emerge, we believe that their most prominent role is to allow neurons to solve difficult problems more efficiently. By efficiency, we refer to the computational cost of solving a problem. We can simplify this to the number of parameters that must be updated and the energy that must be consumed to make all necessary computations and updates. For example, a human neuron can solve XOR using solely its apical dendrites, whereby a nonmonotonic function resides [66]. With point neurons, this would require a small network of linear-non-linear units. Obviously, the biological solution is cheaper and, therefore, more efficient (no need to engage multiple neurons, which have a large volume, and it is hard to activate them nor update their connection weights). Another example comes from associative memory formation: a memory stored in a network of dendritic neurons is encoded by fewer neurons that fire at lower rates, compared to a memory stored in a network of point neurons. At the same time,

dendrites maximize the sparsity of responses and minimize memory interference [27,52]. This is achieved using the exact same number of trainable parameters (synaptic weights) for both types of networks. Overall, if the aforementioned benefits in resource savings could be realized in ML, they would bring major advantages in systems where resource consumption is critical (e.g., in portable devices, mobile phones, microchips, and so forth). Achieving this, however, is not trivial as a one-to-one correspondence between the biological and the ANN components responsible for the observed benefits does not always exist.

Multiscale plasticity rules

The aforementioned savings in biological networks are not only a result of their dendritic structure and active mechanisms. These features act in concert with a set of plasticity rules that dynamically shape not only synaptic weights but also network connectivity and activation functions, and whose contribution to learning is often unintuitive [22]. For example, synaptic plasticity can occur in the absence of somatic spiking [72,73], and synapses can be potentiated or depressed according to the activity of their neighbors [74]. As a result, synapses that fire together are more likely to cluster within the same dendrite, irrespectively of whether they fire strongly enough to excite the cell [21,75,76]. The latter is contrary to the common belief that neuronal firing is necessary for synaptic plasticity (as in classical -but not local- Hebbian rules) and highlights the richness of biological ways to update synaptic weights.

Importantly, the probability of finding co-active synaptic partners is ensured by rewiring, a structural plasticity mechanism that allows synapses to form and retract within dendrites in search of an appropriate partner [77–81]. This type of plasticity is an example of how nature implements dynamics changes in the wiring (connectivity/graph) of biological networks while at the same time facilitating synaptic weight updates. Finally, synapses also undergo homeostatic plasticity (scaling) to prevent run-away excitation (or inhibition) and maintain a high signal-to-noise ratio between potentiated synapses while allowing further plastic changes to take place [82]. This homeostatic plasticity can also be dendrite specific [83].

Beyond synaptic weight and wiring plasticity, the dendrites and somata of biological neurons also undergo plasticity of their intrinsic excitability [23,84,85], which is analogous to changes in their activation functions. For example, learning-induced reductions in A-type K^+ currents can make dendritic activation functions more supralinear and affect the coupling between dendritic and somatic compartments [23,86]. Similarly, changes in the somatic afterhyperpolarization current of neurons involved in learning allow them to become more

excitable and remain at a heightened excitability state for hours [84].

The specific roles subserved by each plasticity type (weight, wiring, activation function) remain unclear, yet some contributions are directly relevant to ML in the sense that they could provide inductive biases. For example, the plasticity of somatic excitability increases the probability that a neuron involved in learning one piece of information will also participate in learning subsequent events (within a window of hours). This property facilitates the linking of information across time, forming memory episodes that extend over several hours [76,87]. Such a learning rule could, for example, empower ANNs that learn sequences (e.g., recurrent ANNs). Moreover, at the single neuron level, temporal binding of information is realized via the co-strengthening (clustering) of synapses from subsequent events in the same dendrites [76]. Synapse clustering results from the combined effects of *rewiring* and *cooperative* plasticity [52,76], it can be established in development and persist through adulthood under a common framework [88], and is a very effective way to bind together pieces of information that arrive simultaneously from different input streams, thus forming associations [10,50]. Such plasticity rules facilitate the detection of regularities in the input distributions sensed by biological networks and bias the system towards learning higher-order correlations. We propose that incorporating rewiring and cooperative plasticity in dendritic ANNs is likely to facilitate learning in dynamically changing environments (such as transfer and continual learning). Recent attempts to implement rewiring in deep ANNs, based on a probabilistic approach, are highly promising with respect to performance accuracy [89].

Finally, another important feature of biological plasticity rules is that they do not need a signal dictating how close the network is to solve a specific problem. Associative learning, for example, can be achieved in a fully unsupervised manner [76]. On the other hand, cortical circuits were recently reported to implement supervised learning via the use of feedback signals [90]. The ability of biological networks to implement both supervised and unsupervised learning is another property that could help empower ANNs. The latter are typically trained using the backpropagation-of-error algorithm (back-prop) although more sophisticated and complex (often bio-inspired) rules have also been applied [91,92]. The back-prop moves along the direction of the minimal gradient, thus rapidly converging to the optimal solution. However, back-prop requires that every synapse in the multiple layers of an ANN knows its contribution to the global error, which is computed at the very last layer. Such a condition may appear biologically implausible, yet various biorealistic solutions have been suggested [93–96]. For example, according to Sacramento et al.

[93], prediction errors are encoded in the distal dendrites of excitatory neurons through the activity of local interneurons and are mediated by local plasticity rules. In support of this idea, Doron and colleagues [90] recently found that the distal dendrites of cortical pyramidal neurons receive such feedback signals that coordinate learning in ways similar to the backpropagation algorithm. Another approach uses error signals that are calculated locally, in a layer-specific manner [97,98], thus eliminating the necessity to propagate the global error back to all hidden layers. The difference target propagation is another alternative [99], whereby targets instead of gradients are computed based on an autoencoding feedback loop, and thus, minimization is achieved with auxiliary variables [100]. In other words, the problem is broken down to easier-to-solve local subproblems via inserting auxiliary variables corresponding to activations in each layer.

Beyond the attempt to find biorealistic solutions to error-backpropagation, other biological learning rules have recently been used in combination with dendritic structure. Combining a neuroinspired plasticity rule based on bursting with a dendritic-like sparse connectivity structure was recently shown to achieve very good classification performance [30]. Moreover, in ANNs with multicompartamental dendrites, an online optimization algorithm was shown to perform canonical correlation analysis, an unsupervised task where the data are projected in a new subspace maximizing their correlations [101]. It should be noted that incorporating more biologically plausible rules in dendritic ANNs does not necessarily improve performance per se, but it often achieves similar performance using fewer resources. For example, according to Illing et al. [102], a local, biologically plausible learning rule based on a layer-specific loss function can learn representations of images and speech with less hardware consumption. Overall, these examples pinpoint the potential benefits of incorporating bioinspired plasticity rules, in conjunction with dendritic structure, to empower ANNs.

Conclusions

This opinion article highlights specific features of biological dendritic trees whose adoption is likely to endow important advantages to ANNs (summarized in Table 1). These features include a) the compartmentalizing and filtering effects of dendritic morphology, which effectively sparsify the connectivity of biological networks and form semi-independent processing units within individual neurons. b) The integrative properties of dendrites, culminating in the generation of dendritic spikes. These spikes transform synaptic inputs in ways similar to those of nonlinear activation functions in artificial nodes while also allowing dendrites to solve difficult spatio-temporal problems, such as logical operations (e.g., AND, OR, XOR, NAND), implement

Table 1

Dendritic features compared to respective approaches in ANNs.

	Artificial neurons in ANNs	Biological neurons
Morphology	Nodes as point neurons, each node can be a multilayered perceptron, symmetric tree	Complex dendritic tree with several branch points, not symmetric
Sparsification	Dropout layers	Inputs are distributed across the tree
Integration	Activation functions used in each node (e.g., sigmoid, tanh, ReLU)	Dendritic spikes (Na^+ , Ca^{2+} , NMDA, dCaAPs, back-propagating action potentials) or passive (sublinear) integration
Learning	Backpropagation-of-error algorithm (most effective)	Cooperative synaptic plasticity (local), plasticity of excitability (dendritic/somatic), bursting
Homeostasis	Regularization of weights/normalization of layers	The total synaptic strength of all synapses (in a branch, neuron, or network) scales up/down
Rewiring	Not typically used, would amount to dynamically changing the graph	Synapses turn over dynamically, and only those that fire with neighbors are stabilized

coincidence and sequence detection, associate input streams, and so forth. c) The dampening effects of morphology and the amplifying effects of dendritic spikes are bound together by a variety of plasticity rules that fine-tune input-output transformations at all levels of processing: from the strength of individual synapses to the grouping of correlated inputs, the coupling strength of each dendrite to its soma, the modulation of excitability and the homeostatic scaling of connections. We believe that these features boost learning and problem-solving in biological brains, not in isolation but through their intricate interplay. They do so by augmenting flexibility so that biological networks can quickly tune into new problems without forgetting the ones learned. This is realized via local updates that are not driven by prediction errors (which are task-specific) but rather by the regularities in the input streams (which can be generic across tasks). Updates occur within specific, sparsely connected compartments, thus allowing different problems to tap onto common features as detected at the bottom layers. This hypothesis is supported by the studies discussed here. Yet, it remains unknown whether dendrites can empower learning in ANNs in a manner that goes well beyond the traditional architectures, enabling them to solve more challenging problems.

What would certainly comprise an important advantage is the physical implementation of dendritic ANNs on neuromorphic chips. Currently, the power consumption of traditional DL architectures is much higher than that of a human brain, while their robustness and flexibility in different tasks remain limited [103]. Neuromorphic computing, on the other hand, can perform complex calculations faster, more power-efficiently, and on a

smaller area than traditional von Neumann architectures [104]. Despite the current advances in the hardware domain, for example, chips stack in a 3D structure [105], the physical dimensions of chips and the need for long wires to emulate the all-to-all connectivity pose severe limitations in implementing state-of-the-art DL architectures on hardware. Dendritic ANNs offer a solution to this problem, as they massively reduce the need for long cables and all-to-all connectivity [106,107]. Recent attempts to integrate dendritic functions on hardware [108] have resulted in a single chip containing the three critical components of a biological neuron; soma, dendrite, and synapses, which can efficiently process spatio-temporal information. These first steps are highly promising towards making neuromorphic computing a practical reality. We hope that the seeds planted here will inspire generations of ML and hardware researchers, helping them achieve major savings in resource utilization and perhaps even human-level performance in tasks that remain unsolvable by current systems.

Author contributions

S.C. and P.P. conceived and wrote the paper.

Conflict of interest statement

Nothing declared.

Acknowledgements

The authors thank Blake Richards, Konrad Kording, and Richard Naud for their thoughtful comments and efforts towards improving our manuscript. This work was supported by the NEUREKA grant (H2020-FETOPEN-2018-2019-2020-01, FET-Open Challenging Current Thinking, GA-863245) and the Einstein Foundation Berlin, Germany (visiting fellowship to P.P.).

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest

** of outstanding interest

1. LeCun Y, Bengio Y, Hinton G: **Deep learning**. *Nature* 2015, **521**: 436–444.
2. Nassif AB, Shahin I, Attili I, Azzeh M, Shaalan K: **Speech recognition using deep neural networks: a systematic review**. *IEEE Access* 2019, **7**:19143–19165.
3. Masi I, Wu Y, Hassner T, Natarajan P: **Deep face recognition: a survey**. In 2018 31st SIBGRAPI Conference on graphics, Patterns and images (SIBGRAPI). IEEE; 2018:471–478.
4. Justesen N, Bontrager P, Togelius J, Risi S: **Deep learning for video game playing**. *IEEE Trans Games* 2020, **12**:1–20.
5. Huang Y, Chen Y: **Survey of state-of-art autonomous driving technologies with deep learning**. In *IEEE 20th international conference on software quality, reliability and security companion (QRS-C)*. IEEE; 2020:221–228.
6. Stuart G, Spruston N, Häusser M: *Dendrites*. Oxford University Press; 2016.
7. Spruston N, Jaffe DB, Johnston D: **Dendritic attenuation of synaptic potentials and currents: the role of passive membrane properties**. *Trends Neurosci* 1994, **17**:161–166.
8. Major G, Larkum ME, Schiller J: **Active properties of neocortical pyramidal neuron dendrites**. *Annu Rev Neurosci* 2013, **36**:1–24.
9. Softky W: **Sub-millisecond coincidence detection in active dendritic trees**. *Neuroscience* 1994, **58**:13–41.
10. Ariav G, Polsky A, Schiller J: **Submillisecond precision of the input-output transformation function mediated by fast sodium dendritic spikes in basal dendrites of CA1 pyramidal neurons**. *J Neurosci* 2003, **23**:7750–7758.
11. Spruston N: **Pyramidal neurons: dendritic structure and synaptic integration**. *Nat Rev Neurosci* 2008, **9**:206–221.
12. Ujfalussy BB, Makara JK, Lengyel M, Branco T: **Global and multiplexed dendritic computations under in vivo-like conditions**. *Neuron* 2018, **100**:579–592.e5.
13. Polsky A, Mel BW, Schiller J: **Computational subunits in thin dendrites of pyramidal cells**. *Nat Neurosci* 2004, **7**:621–627.
14. Branco T, Häusser M: **Synaptic integration gradients in single cortical pyramidal cell dendrites**. *Neuron* 2011, **69**:885–892.
15. Ujfalussy B, Kiss T, Érdi P: **Parallel computational subunits in dentate granule cells generate multiple place fields**. *PLoS Comput Biol* 2009, **5**, e1000500.
16. Bono J, Wilmes KA, Clopath C: **Modelling plasticity in dendrites: from single cells to networks**. *Curr Opin Neurobiol* 2017, **46**:136–141.
17. Kastellakis G, Poirazi P: **Synaptic clustering and memory formation**. *Front Mol Neurosci* 2019, **12**.
18. Gordon U, Polsky A, Schiller J: **Plasticity compartments in basal dendrites of neocortical pyramidal neurons**. *J Neurosci* 2006, **26**:12717–12726.
19. Sajikumar S, Korte M: **Different compartments of apical CA1 dendrites have different plasticity thresholds for expressing synaptic tagging and capture**. *Learn Mem* 2011, **18**:327–331.
20. Mel BW, Schiller J, Poirazi P: **Synaptic plasticity in dendrites: complications and coping strategies**. *Curr Opin Neurobiol* 2017, **43**:177–186.
21. Weber JP, Andrásfalvy BK, Polito M, Magó Á, Ujfalussy BB, Makara JK: **Location-dependent synaptic plasticity rules by dendritic spine cooperativity**. *Nat Commun* 2016, **7**:11380.
22. Magee JC, Grienberger C: **Synaptic plasticity forms and functions**. *Annu Rev Neurosci* 2020, **43**:95–117.
23. Losonczy A, Makara JK, Magee JC: **Compartmentalized dendritic plasticity and input feature storage in neurons**. *Nature* 2008, **452**:436–441.
24. Poirazi P, Brannon T, Mel BW: **Pyramidal neuron as two-layer neural network**. *Neuron* 2003, **37**:989–999.
25. Jadi M, Behabadi BF, Poleg-Polsky A, Schiller J, Mel BW: **An augmented two-layer model captures nonlinear analog spatial integration effects in pyramidal neuron dendrites**. *Proc IEEE* 2014, **102**:782–798.
26. Mlil M, London M, Häusser M: **Dendritic computation**. *Annu Rev Neurosci* 2005, **28**:503–532.
27. Tzivilaki A, Kastellakis G, Poirazi P: **Challenging the point neuron dogma: FS basket cells as 2-stage nonlinear integrators**. *Nat Commun* 2019, **10**:3664.
28. Branco T, Häusser M: **The single dendritic branch as a fundamental functional unit in the nervous system**. *Curr Opin Neurobiol* 2010, **20**:494–502.
29. Guerguiev J, Lillicrap TP, Richards BA: **Towards deep learning with segregated dendrites**. *Elife* 2017, **6**, e22901.
30. Payeur A, Guerguiev J, Zenke F, Richards BA, Naud R: **Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits**. *bioRxiv* 2020, <https://doi.org/10.1101/2020.03.30.015511>.
31. Siegel M, Körding KP, König P: **Integrating top-down and bottom-up sensory processing by somato-dendritic interactions**. *J Comput Neurosci* 2000, **8**:161–173.
32. Kreiman G, Serre T: **Beyond the feedforward sweep: feedback computations in the visual cortex**. *Ann N Y Acad Sci* 2020, **1464**:222–241.
33. Körding KP, König P: **Supervised and unsupervised learning with two sites of synaptic integration**. *J Comput Neurosci* 2001, **11**:207–215.
34. Kirkpatrick J, Pascanu R, Rabinowitz N, Veness J, Desjardins G, Rusu AA, Milan K, Quan J, Ramalho T, Grabska-Barwinska A, et al.: **Overcoming catastrophic forgetting in neural networks**. *Proc Natl Acad Sci U S A* 2017, **114**:3521–3526.
35. Zador AM, Agmon-Snir H, Segev I: **The morphoelectronic transform: a graphical approach to dendritic function**. *J Neurosci* 1995, **15**:1669–1682.
36. Stuart G, Spruston N: **Determinants of voltage attenuation in neocortical pyramidal neuron dendrites**. *J Neurosci* 1998, **18**:3501–3510.
37. Baruah SMB, Nandi D, Roy S: **Modelling signal transmission in passive dendritic fibre using discretized cable equation**. *Proc 2nd Int Conf Innov Electron Signal Process Commun IESC* 2019, <https://doi.org/10.1109/IESPC.2019.8902444>.
38. Ferrante M, Migliore M, Ascoli GA: **Functional impact of dendritic branch-point morphology**. *J Neurosci* 2013, **33**:2156–2165.
39. Rall W: **Theoretical significance of dendritic trees for neuronal input-output relations**. In *Neural theory and modeling*. Edited by Reiss R, Stanford UP; 1964:63–97.
40. Cazé RD, Humphries M, Gutkin B: **Passive dendrites enable single neurons to compute linearly non-separable functions**. *PLoS Comput Biol* 2013, **9**, e1002867.
41. Brunel N, Hakim V, Isope P, Nadal J-P, Barbours B: **Optimal information storage and the distribution of synaptic weights**. *Neuron* 2004, **43**:745–757.
42. Chavlis S, Petrantonakis PC, Poirazi P: **Dendrites of dentate gyrus granule cells contribute to pattern separation by controlling sparsity**. *Hippocampus* 2017, **27**:89–110.
43. Schweighofer N, Doya K, Lay F: **Unsupervised learning of granule cell sparse codes enhances cerebellar adaptive control**. *Neuroscience* 2001, **103**:35–50.

44. Hawkins J, Ahmad S: **Why neurons have thousands of synapses, a theory of sequence memory in neocortex.** *Front Neural Circ* 2016, **10**.
 45. McCulloch WS, Pitts W: **A logical calculus of the ideas immanent in nervous activity.** *Bull Math Biol* 1990, **52**:99–115.
 46. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R: **Dropout: a simple way to prevent neural networks from over-fitting.** *J Mach Learn Res* 2014, **15**:1929–1958.
 47. Shukla A, Tiwari R, Kala R: **Modular neural networks.** In *Towards hybrid and adaptive computing. Studies in computational intelligence.* Springer; 2010:307–335.
 48. Poirazi P, Neocleous C, Pattichis CS, Schizas CN: **Classification capacity of a modular neural network implementing neurally inspired architecture and training.** *IEEE Trans Neural Networks* 2004, **18**:597–612.
 49. Liu Baoyuan, Wang Min, Foroosh H, Tappen M, Pensky M: **Sparse convolutional neural networks.** In *IEEE conference on computer vision and pattern recognition (CVPR).* IEEE; 2015: 806–814.
 50. Losonczy A, Magee JC: **Integrative properties of radial oblique dendrites in hippocampal CA1 pyramidal neurons.** *Neuron* 2006, **50**:291–307.
 51. Poirazi P, Brannon T, Mel BW: **Arithmetic of subthreshold synaptic summation in a model CA1 pyramidal cell.** *Neuron* 2003, **37**:977–987.
 52. Poirazi P, Mel BW: **Impact of active dendrites and structural plasticity on the memory capacity of neural tissue.** *Neuron* 2001, **29**:779–796.
 53. Golding NL, Spruston N: **Dendritic sodium spikes are variable triggers of axonal action potentials in hippocampal CA1 pyramidal neurons.** *Neuron* 1998, **21**:1189–1200.
 54. Milojkovic BA, Wuskell JP, Loew LM, Antic SD: **Initiation of sodium spikelets in basal dendrites of neocortical pyramidal neurons.** *J Membr Biol* 2005, **208**:155–169.
 55. Larkum ME, Nevian T, Sandler M, Polsky A, Schiller J: **Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle.** *Science (80-)* 2009, **325**:756–760.
 56. Schiller J, Schiller Y, Stuart G, Sakmann B: **Calcium action potentials restricted to distal apical dendrites of rat neocortical pyramidal neurons.** *J Physiol* 1997, **505**:605–616.
 57. Cichon J, Gan W-B: **Branch-specific dendritic Ca²⁺ spikes cause persistent synaptic plasticity.** *Nature* 2015, **520**:180–185.
 58. Brunner J, Szabadics J: **Analogue modulation of back-propagating action potentials enables dendritic hybrid signalling.** *Nat Commun* 2016, **7**:13033.
 59. Silver RA: **Neuronal arithmetic.** *Nat Rev Neurosci* 2010, **11**: 474–489.
 60. Larkum ME, Zhu JJ, Sakmann B: **A new cellular mechanism for coupling inputs arriving at different cortical layers.** *Nature* 1999, **398**:338–341.
 61. Beniaquev D, Segev I, London M: **Single cortical neurons as deep artificial neural networks.** *bioRxiv* 2020, <https://doi.org/10.1101/613141>.
 62. Häusser M, Mel B: **Dendrites: bug or feature?** *Curr Opin Neurobiol* 2003, **13**:372–383.
 63. Chiovini B, Turi GF, Katona G, Kaszás A, Pálfi D, Maák P, Szalay G, Szabó MF, Szadai Z, et al.: **Dendritic spikes induce ripples in parvalbumin interneurons during hippocampal sharp waves.** *Neuron* 2014, **82**:908–924.
 64. Katona G, Kaszás A, Turi GF, Hájos N, Tamás G, Vizi ES, Rózsa B: **Roller Coaster Scanning reveals spontaneous triggering of dendritic spikes in CA1 interneurons.** *Proc Natl Acad Sci* 2011, **108**:2148–2153.
 65. Tran-Van-Minh A, Abrahamsson T, Cathala L, DiGregorio DA: **Differential dendritic integration of synaptic potentials and calcium in cerebellar interneurons.** *Neuron* 2016, **91**:837–850.
 66. Gidon A, Zolnik TAA, Fidzinski P, Bolduan F, Papoutsis A, Poirazi P, Holtkamp M, Vida I, Larkum MEE: **Dendritic action potentials and computation in human layer 2/3 cortical neurons.** *Science* 2020, **367**:83–87.
- This study discovered a novel type of dendritic spikes in the human layer 2/3 pyramidal neurons, whose amplitude decreases with the stimulus intensity. These dendritic calcium-activated spikes enable human neurons to solve the exclusive OR problem.
67. Hayou S, Doucet A, Rousseau J: **On the impact of the activation function on deep neural networks training.** In *Proceedings of the 36th international conference on machine learning*; 2019.
 68. Wu X, Liu X, Li W, Wu Q: **Improved expressivity through dendritic neural networks.** In *Advances in neural information processing systems*; 2018:8057–8068.
 69. Jones IS, Kording KP: **Can single neurons solve MNIST? The computational power of biological dendritic trees.** *arXiv* 2020.
 70. Camp B, Mandivarapu JK, Estrada R: **Continual learning with deep artificial neurons.** *arXiv* 2020.
 71. Georgescu MI, Ionescu RT, Ristea NC, Sebe N: **Non-linear neurons with human-like apical dendrite activations.** *arXiv* 2020, <https://doi.org/10.36227/techrxiv.11830761>.
 72. Bittner KC, Milstein AD, Grienberger C, Romani S, Magee JC: **Behavioral time scale synaptic plasticity underlies CA1 place fields.** *Science (80-)* 2017, **357**:1033–1036.
- This paper shows that place-fields in CA1 pyramidal neurons can emerge as a result of synaptic plasticity that depends on the generation of dendritic plateau potentials. The novel plasticity rule operates in the order of seconds, that is, inputs that occur seconds before or after a dendritic plateau potential are potentiated, leading to the creation of a place field at the location of those inputs.
73. Kampa BM, Letzkus JJ, Stuart GJ: **Dendritic mechanisms controlling spike-timing-dependent synaptic plasticity.** *Trends Neurosci* 2007, **30**:456–463.
 74. El-Boustani S, Ip JPK, Breton-Provencher V, Knott GW, Okuno H, Bito H, Sur M: **Locally coordinated synaptic plasticity of visual cortex neurons in vivo.** *Science (80-)* 2018, **360**:1349–1354.
 75. Lee KFH, Soares C, Thivierge JP, Béique JC: **Correlated synaptic inputs drive dendritic calcium amplification and cooperative plasticity during clustered synapse development.** *Neuron* 2016, **89**:784–799.
 76. Kastellakis G, Silva AJ, Poirazi P: **Linking memories across time via neuronal and dendritic overlaps in model neurons with active dendrites.** *Cell Rep* 2016, **17**:1491–1504.
 77. DeBello W, Zito K: **Within a Spine's reach.** In *The rewiring brain: a computational approach to structural plasticity in the adult brain*; 2017.
 78. DeBello WM: **Micro-rewiring as a substrate for learning.** *Trends Neurosci* 2008, **31**:577–584.
 79. Chklovskii D, Mel BW, Svoboda K: **Cortical rewiring and information storage.** *Nature* 2004, **431**:782–788.
 80. Legenstein R, Maass W: **Branch-specific plasticity enables self-organization of nonlinear computation in single neurons.** *J Neurosci* 2011, **31**:10787–10802.
 81. Limbacher T, Legenstein R: **Emergence of stable synaptic clusters on dendrites through synaptic rewiring.** *Front Comput Neurosci* 2020:14.
 82. Turrigiano GG: **The self-tuning neuron: synaptic scaling of excitatory synapses.** *Cell* 2008, **135**:422–435.
 83. Rabinowitch I, Segev I: **The interplay between homeostatic synaptic plasticity and functional dendritic compartments.** *J Neurophysiol* 2006, **96**:276–283.
 84. Disterhoft JF, Oh MM: **Learning, aging and intrinsic neuronal plasticity.** *Trends Neurosci* 2006, **29**:587–599.
 85. Ohtsuki G, Piochon C, Adelman JP, Hansel C: **SK2 channel modulation contributes to compartment-specific dendritic plasticity in cerebellar purkinje cells.** *Neuron* 2012, **75**: 108–120.

86. Makara JK, Losonczy A, Wen Q, Magee JC: **Experience-dependent compartmentalized dendritic plasticity in rat hippocampal CA1 pyramidal neurons.** *Nat Neurosci* 2009, **12**: 1485–1487.
87. Cai DJ, Aharoni D, Shuman T, Shobe J, Biane J, Song W, Wei B, Veshkini M, La-Vu M, Lou J, *et al.*: **A shared neural ensemble links distinct contextual memories encoded close in time.** *Nature* 2016, **534**:115–118.
88. Kirchner JH, Gjorgjieva J: **A unifying framework for synaptic organization on cortical dendrites.** *bioRxiv* 2019, <https://doi.org/10.1101/771907>.
89. Bellec G, Kappel D, Maass W, Legenstein R: **Deep Rewiring: training very sparse deep networks.** In *6th international conference on learning representations*; 2018.
90. Doron G, Shin JN, Takahashi N, Drüke M, Bocklisch C, Skenderi S, De Mont L, Toumazou M, Ledderose J, Brecht M, *et al.*: **Perirhinal input to neocortical layer 1 controls learning.** *Science (80-)* 2020, **370**, eaaz3136.
 This study demonstrates that neurons located in the deep layers of the perirhinal cortex exhibit increased firing after microstimulation upon learning. Learning is associated with the emergence of a small population of neurons in layer 5 of the somatosensory cortex that increased bursting upon stimulation. This increase in bursting was accompanied by an increase in dendritic activity, and silencing the perirhinal cortex to layer 1 projection effectively disrupted learning and its physiological correlates.
91. Illing B, Gerstner W, Brea J: **Biologically plausible deep learning — But how far can we go with shallow networks?** *Neural Networks* 2019, **118**:90–101.
92. Pérez-Sánchez B, Fontenla-Romero O, Guijarro-Berdiñas B: **A review of adaptive online learning for artificial neural networks.** *Artif Intell Rev* 2018, **49**:281–299.
93. Sacramento J, Bengio Y, Costa RP, Senn W: **Dendritic cortical microcircuits approximate the backpropagation algorithm.** In *32nd conference on neural information processing systems*; 2018:8721–8732.
 This paper introduces a computational model whereby distal dendrites encode error signals with the help of dendritic-targeted interneurons. This work is one of the first attempts to explain how the brain solves the credit assignment problem.
94. Lillicrap TP, Cownden D, Tweed DB, Akerman CJ: **Random synaptic feedback weights support error backpropagation for deep learning.** *Nat Commun* 2016, **7**:13276.
 This paper demonstrates that feedback pathways should not necessarily be symmetric to feedforward pathways in ANNs. Thus, the weights used for backpropagation are not the same as the ones used in forward propagation, rather they are randomly chosen.
95. Richards BA, Lillicrap TP: **Dendritic solutions to the credit assignment problem.** *Curr Opin Neurobiol* 2019, **54**:28–36.
96. Lillicrap TP, Santoro A, Marris L, Akerman CJ, Hinton G: **Back-propagation and the brain.** *Nat Rev Neurosci* 2020, **21**:335–346.
97. Mostafa H, Ramesh V, Cauwenberghs G: **Deep supervised learning using local errors.** *Front Neurosci* 2018, **12**:608.
98. Nøkland A, Eidnes LH: **Training neural networks with local error signals.** *Proc 36th Int Conf Mach Learn* 2019, **97**:4839–4850.
 This work implements a layer-specific training of the hidden nodes using locally generated errors. Thus, the weights do not depend on a global error, and can be updated during the forward pass.
99. Lee D-H, Zhang S, Fischer A, Bengio Y: **Difference target propagation.** In *Machine learning and knowledge discovery in databases*; 2015:498–515.
100. Choromanska A, Cowen B, Kumaravel S, Luss R, Rigotti M, Rish I, Diachille P, Gurev V, Kingsbury B, Tejwani R, *et al.*: **Beyond backprop: online alternating minimization with auxiliary variables.** In *Proceedings of the 36th international conference on machine learning*; 2019.
101. Lipshutz D, Bahroun Y, Golkar S, Sengupta AM, Chkrovskii DB: **A biologically plausible neural network for multi-channel Canonical Correlation Analysis.** *arXiv* 2020.
102. Illing B, Gerstner W, Bellec G: **Towards truly local gradients with CLAPP: contrastive, local and predictive plasticity.** *arXiv* 2020.
103. Tang J, Yuan F, Shen X, Wang Z, Rao M, He Y, Sun Y, Li X, Zhang W, Li Y, *et al.*: **Bridging biological and artificial neural networks with emerging neuromorphic devices: fundamentals, progress, and challenges.** *Adv Mater* 2019, **31**.
104. Marković D, Mizrahi A, Querlioz D, Grollier J: **Physics for neuromorphic computing.** *Nat Rev Phys* 2020, **2**:499–510.
105. Lin P, Li C, Wang Z, Li Y, Jiang H, Song W, Rao M, Zhuo Y, Upadhyay NK, Barnell M, *et al.*: **Three-dimensional memristor circuits as complex neural networks.** *Nat Electron* 2020, **3**: 225–232.
106. Roy S, Banerjee A, Basu A: **Liquid state machine with dendritically enhanced readout for low-power, neuromorphic VLSI implementations.** *IEEE Trans Biomed Circuits Syst* 2014, **8**:681–695.
107. Hussain S, Liu S-C, Basu A: **Hardware-amenable structural learning for spike-based pattern classification using a simple model of active dendrites.** *Neural Comput* 2015, **27**:845–897.
108. Li X, Tang J, Zhang Q, Gao B, Yang JJ, Song S, Wu W, Zhang W, Yao P, Deng N, *et al.*: **Power-efficient neural network with artificial dendrites.** *Nat Nanotechnol* 2020, **15**:776–782.
 This paper introduces the first neuromorphic chip with artificial, functional dendrites. The network shows promising performance against noisy datasets with a total reduction of the energy consumption.