2021 Special Issue on AI and Brain Science: Perspective

Natural and Artificial Intelligence: A brief introduction to the interplay between AI and neuroscience research

Tom Macpherson [a], Anne Churchland [b], Terry Sejnowski [c,d], James DiCarlo [e], Yukiyasu Kamitani [f,g], Hidehiko Takahashi [h], Takatoshi Hikida [a,*]

[a] *Laboratory for Advanced Brain Functions, Institute for Protein Research, Osaka University, Osaka, Japan*
[b] *Cold Spring Harbor Laboratory, Neuroscience, Cold Spring Harbor, NY, USA*
[c] *Computational Neurobiology Laboratory, Salk Institute for Biological Studies, CA, USA*
[d] *Division of Biological Sciences, University of California San Diego, CA, USA*
[e] *Brain and Cognitive Sciences, Massachusetts Institute of Technology, MA, USA*
[f] *Department of Neuroinformatics, ATR Computational Neuroscience Laboratories, Kyoto, Japan*
[g] *Graduate School of Informatics, Kyoto University, Kyoto, Japan*
[h] *Department of Psychiatry and Behavioral Sciences, Tokyo Medical and Dental University Graduate School, Tokyo, Japan*

ARTICLE INFO

ABSTRACT

Neuroscience and artificial intelligence (AI) share a long history of collaboration. Advances in neuroscience, alongside huge leaps in computer processing power over the last few decades, have given rise to a new generation of *in silico* neural networks inspired by the architecture of the brain. These AI systems are now capable of many of the advanced perceptual and cognitive abilities of biological systems, including object recognition and decision making. Moreover, AI is now increasingly being employed as a tool for neuroscience research and is transforming our understanding of brain functions. In particular, deep learning has been used to model how convolutional layers and recurrent connections in the brain's cerebral cortex control important functions, including visual processing, memory, and motor control. Excitingly, the use of neuroscience-inspired AI also holds great promise for understanding how changes in brain networks result in psychopathologies, and could even be utilized in treatment regimes. Here we discuss recent advancements in four areas in which the relationship between neuroscience and AI has led to major advancements in the field; (1) AI models of working memory, (2) AI visual processing, (3) AI analysis of big neuroscience datasets, and (4) computational psychiatry.

1. Introduction

Classically, our definition of intelligence has largely been based upon the capabilities of advanced biological entities, most notably humans. Accordingly, research into artificial intelligence (AI) has primarily focused on the creation of machines that can perceive, learn, and reason, with the overarching objective of creating an artificial general intelligence (AGI) system that can emulate human intelligence, so called Turing-powerful systems. Considering this aim, it is not surprising that scientists, mathematicians, and philosophers working on AI have taken inspiration from the mechanistic, structural, and functional properties of the brain.

Since at least the 1950s, attempts have been made to artificially model the information processing mechanisms of neurons. This primarily began with the development of perceptrons (Rosenblatt, 1958), a highly reductionist model of neuronal signaling, in which an individual node receiving weighted inputs could produce a binary output if the summation of inputs reached a threshold. Coinciding with the emergence of the cognitive revolution in the 50s and 60s, and extending until at least the 90s, there was initially much pushback against the development of artificial neural networks within the AI and cognitive science communities (Fodor & Pylyshyn, 1988; Mandler, 2002; Minsky & Papert, 1969). However, by the late 1980s, the development of multilayer neural networks and the popularization of backpropagation had solved many of the limitations of early perceptrons, including their inability to solve non-linear classification problems such as learning a simple boolean XOR function (Rumelhart, Hinton, & Williams, 1986). Neural networks were now able to dynamically modify their own connections by calculating error functions of the network and communicating them back through

* Correspondence to: Laboratory for Advanced Brain Functions, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan.
E-mail address: hikida@protein.osaka-u.ac.jp (T. Hikida).

the constituent layers, giving rise to a new generation of AI capable of intelligent skills including image and speech recognition (Bengio, 1993; LeCun, Bengio, & Hinton, 2015; LeCun et al., 1989). To date, backpropagation is still commonly used to train deep neural networks (Lillicrap, Santoro, Marris, Akerman, & Hinton, 2020; Richards et al., 2019), and has been combined with reinforcement learning methods to create advanced learning systems capable of matching or outperforming humans in strategy-based games including Chess (Silver et al., 2018), Go (Silver et al., 2016, 2018), poker (Moravčík et al., 2017), and StarCraft II (Vinyals et al., 2019).

Despite the neuroscience-inspired origins of AI, the biological plausibility of modern AI is questionable. Indeed, there is little evidence that backpropagation of error underlies the modification of synaptic connections between neurons (Crick, 1989; Grossberg, 1987); although recent theories have suggested that an approximation of backpropagation may exist in the brain (Lillicrap et al., 2020; Whittington & Bogacz, 2019). While creating brain-like systems is clearly not necessary to achieve all goals of AI, as evidenced by the above-described accomplishments, a major advantage of biologically plausible AI is its usefulness for understanding and modeling information processing in the brain. Additionally, brain mechanisms can be thought of as an evolutionarily-validated template for intelligence, honed over millions of years for adaptability, speed, and energy efficiency. As such, increased integration of brain-inspired mechanisms may help to further improve the capabilities and efficiency of AI. These ideas have led to continued interest in the creation of neuroscience-inspired AI, and have further strengthened the partnership between AI and neuroscience research fields. In the last decade, several biologically plausible alternatives to backpropagation have been suggested, including predictive coding (Bastos et al., 2012; Millidge, Tschantz, & Buckley, 2020), feedback alignment (Lillicrap, Cownden, Tweed, & Akerman, 2016), equilibrium propagation (Scellier & Bengio, 2017), Hebbian-like learning rules (Krotov & Hopfield, 2019; Miconi, 2017), and zero-divergence inference learning (Salvatori, Song, Lukasiewicz, Bogacz, & Xu, 2021). Similarly, other recent efforts to bridge the gap between artificial and biological neural networks have led to the development of spiking neural networks capable of approximating stochastic potential-based communication between neurons (Pfeiffer & Pfeil, 2018), as well as attention-like mechanisms including transformer architectures (Vaswani et al., 2017).

The beneficial relationship between AI and neuroscience is reciprocal, and AI is now rapidly becoming an invaluable tool in neuroscience research. AI models designed to perform intelligence-based tasks are providing novel hypotheses for how the same processes are controlled within the brain. For example, work on distributional reinforcement learning in AI has recently resulted in the proposal of a new theory of dopaminergic signaling of probabilistic distributions (Dabney et al., 2020). Similarly, goal-driven deep learning models of visual processing have been used to estimate the organizational properties of the brain's visual system and accurately predict patterns of neural activity (Yamins & DiCarlo, 2016). Additionally, advances in deep learning algorithms and the processing power of computers now allow for high-throughput analysis of large-scale datasets, including that of whole-brain imaging in animals and humans, expediting the progress of neuroscience research (Thomas, Heekeren, Müller, & Samek, 2019; Todorov et al., 2020; Zhu et al., 2019). Deep learning models trained to decode neural imaging data can create accurate predictions of decision-making, action selection, and behavior, helping us to understand the functional role of neural activity, a key goal of cognitive neuroscience (Batty et al., 2019; Musall, Urai, Sussillo, & Churchland, 2019). Excitingly, machine learning and deep learning approaches are also now being applied to the

emerging field of computational psychiatry to simulate normal and dysfunctional brain states, as well as to identify aberrant patterns of brain activity that could be used as robust classifiers for brain disorders (Cho, Yim, Choi, Ko, & Lee, 2019; Durstewitz, Koppe, & Meyer-Lindenberg, 2019; Koppe, Meyer-Lindenberg, & Durstewitz, 2021; Zhou et al., 2020).

In the last few years, several reviews have examined the long and complicated relationship between neuroscience and AI (see Hassabis, Kumaran, Summerfield, & Botvinick, 2017; Hasson, Nastase, & Goldstein, 2020; Kriegeskorte & Douglas, 2018; Richards et al., 2019; Ullman, 2019). Here, we aim to give a brief introduction into how the interplay between neuroscience and AI fields has stimulated progress in both areas, focusing on four important themes taken from talks presented at a symposium entitled "AI for neuroscience and neuromorphic technologies" at the 2020 International Symposium on Artificial Intelligence and Brain Science; (1) AI models of working memory, (2) AI visual processing, (3) AI analysis of neuroscience data, and (4) computational psychiatry. Specifically, we focus on how recent neuroscience-inspired approaches are resulting in AI that is increasingly brain-like and is not only able to achieve human-like feats of intelligence, but is also capable of decoding neural activity and accurately predicting behavior and the brain's mental contents. This includes spiking and recurrent neural network models of working memory inspired by the stochastic spiking properties of biological neurons and their sustained activation during memory retention, as well as neural network models of visual processing incorporating convolutional layers inspired by the architecture of the brain's visual ventral stream. Additionally, we discuss how AI is becoming an increasingly powerful tool for neuroscientists and clinicians, acting as diagnostic and even therapeutic aids, as well as potentially informing us about brain mechanisms, including information processing and memory.

2. Neuroscience-inspired artificial working memory

One of the major obstacles of creating brain-like AI systems has been the challenge of modeling working memory, an important component of intelligence. Today, most *in silico* systems utilize a form of working memory known as random-access memory (RAM) that acts as a cache for data required by the central processor and is separated from long-term memory storage in solid state or hard disk drives. However, this architecture differs considerably from the brain, where working and long-term memory appear to involve, at least partly, the same neural substrates, predominantly the neocortex (Baddeley, 2003; Blumenfeld & Ranganath, 2007; Rumelhart & McClelland, 1986; Shimamura, 1995) and hippocampus (Bird & Burgess, 2008; Eichenbaum, 2017). These findings suggest that within these regions, working memory is likely realized by specific brain mechanisms that allow for fast and short-term access to information.

Working memory tasks performed in humans and non-human primates have indicated that elevated and persistent activity within cell assemblies of the prefrontal cortex, as well as other areas of the neocortex, hippocampus, and brainstem, may be critical for information retention within working memory (Boran et al., 2019; Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017; Fuster & Alexander, 1971; Goldman-Rakic, 1995; McFarland & Fuchs, 1992; Miller, Erickson, & Desimone, 1996; Watanabe & Niki, 1985). In response to these findings, several neural mechanisms have been proposed to account for this persistent activity (reviewed in Durstewitz, Seamans, & Sejnowski, 2000). These include recurrent excitatory connectivity between networks of neurons (Hopfield, 1982; O'Reilly, Braver, & Cohen, 1999), *cellular bistability*, where the intrinsic properties of neurons can produce a continuously spiking state (Lisman, Fellous,

& Wang, 1998; Marder, Abbott, Turrigiano, Liu, & Golowasch, 1996; O'Reilly et al., 1999), and *synfire chains*, where activity is maintained in synchronously firing feed-forward loops (Diesmann, Gewaltig, & Aertsen, 1999; Prut et al., 1998). Of these, the most widely researched have been models of persistent excitation in recurrently connected neural networks. These began with simple networks, such as recurrent attractor networks, where discrete working memories represent the activation of attractors, stable patterns of activity in networks of neurons reciprocally connected by strong synaptic weights formed by Hebbian learning (Amit, Bernacchia, & Yakovlev, 2003; Amit & Brunel, 1995; Durstewitz et al., 2000). Afferent input to these networks strong enough to reach a threshold will trigger recurrent excitation and induce a suprathreshold of excitation that persists even when the stimulus is removed, maintaining the stimulus in working memory. Subsequent and more complex computational models have demonstrated recurrent networks connecting the cortex, basal ganglia, and thalamus to be capable of working memory maintenance, and to be able to explain patterns of neural activity observed in neurophysiological studies of working memory (Beiser & Houk, 1998; Botvinick & Plaut, 2006; Hazy, Frank, & O'Reilly, 2007; O'Reilly et al., 1999; Zipser, 1991).

Inspired by above-described biological and computational studies, artificial recurrent neural networks (RNNs) have been designed as a model of the recurrent connections between neurons within the brain's cerebral cortex. These RNNs have since been reported to be capable of performing a wide variety of cognitive tasks requiring working memory (Botvinick & Plaut, 2006; Mante, Sussillo, Shenoy, & Newsome, 2013; Rajan, Harvey, & Tank, 2016; Song, Yang, & Wang, 2016; Sussillo & Abbott, 2009; Yang, Joglekar, Song, Newsome, & Wang, 2019). More recently, researchers have been working on a new generation of spiking recurrent neural networks (SRNN), aiming to recreate the stochastic spiking properties of biological circuits and demonstrating similar performance in cognitive tasks to the above-described continuous-rate RNN models (Kim, Li, & Sejnowski, 2019; Xue, Halassa, & Chen, 2021; Yin, Corradi, & Bohté, 2020). These spiking networks not only aim to achieve greater energy efficiency, but also provide improved biological plausibility, offering advantages for modeling and potentially informing how working memory may be controlled in the brain (Diehl, Zarrella, Cassidy, Pedroni, & Neftci, 2016; Han, Sengupta, & Roy, 2016; Pfeiffer & Pfeil, 2018; Taherkhani et al., 2020). Indeed, in a recent study, an SRNN trained on a working memory task was revealed to show remarkably similar temporal properties to single neurons in the primate PFC (Kim & Sejnowski, 2021). Further analysis of the model uncovered the existence of a disinhibitory microcircuit that acts as a critical component for long neuronal timescales that have previously been implicated in working memory maintenance in real and simulated networks (Chaudhuri, Knoblauch, Gariel, Kennedy, & Wang, 2015; Wasmuht, Spaak, Buschman, Miller, & Stokes, 2018). The authors speculate that recurrent networks with similar inhibitory microcircuits may be a common feature of cortical regions requiring short-term memory maintenance, suggesting an interesting avenue of study for neuroscientists researching working memory mechanisms in the brain.

Finally, it is important to note that while there is clear biological evidence for spiking activity during memory retention periods in working memory tasks, the majority of studies reporting persistent activity during these periods calculated the averaged spiking activity across trials, potentially masking important intra-trial spiking dynamics (Lundqvist, Herman, & Miller, 2018). Interestingly, recent single-trial analyses of working memory tasks suggests that frontal cortex networks demonstrate sparse, transient coordinated bursts of spiking activity, rather than persistent activation (Bastos, Loonis, Kornblith, Lundqvist, & Miller, 2018;

Lundqvist, Herman, Warden, Brincat, & Miller, 2018; Lundqvist et al., 2016). Such patterns of neural activity may be explained by models of transient spiking activity, such as the "synaptic attractor" model, where working memories are maintained by spike-induced Hebbian synaptic plasticity in between transient coordinated bursts of activity (Fiebig & Lansner, 2017; Huang & Wei, 2021; Lundqvist, Herman, & Miller, 2018; Mongillo, Barak, & Tsodyks, 2008; Sandberg, Tegnér, & Lansner, 2003). These models suggest that synaptic plasticity may allow working memory to be temporarily stored in an energy efficient manner that is also less susceptible to interference, while bursts of spiking may allow for fast reading of information when necessary (Huang & Wei, 2021; Lundqvist, Herman, & Miller, 2018). Further investigation of working memory in biological studies using single-trial analyses, as well as neuroscience-inspired AI models trained on working memory tasks, may help to elucidate precisely when and how these spiking and plasticity-based processes are utilized within the brain.

Here we have discussed how neuroscience findings over the last few decades inspired the creation of computational models of working memory in humans and non-human primates. These studies subsequently informed the creation of artificial neural networks designed to model the organization and function of brain networks, including the inclusion of recurrent connections between neurons and the introduction of spiking properties. The relationship between neuroscience and AI research has now come full circle, with recent SRNN models potentially informing about brain mechanisms underlying working memory (Kim & Sejnowski, 2021). In the next section we continue this examination of the benefits of the partnership between neuroscience and AI research, discussing how brain architectures have inspired the design of artificial visual processing models and how brain imaging data has been used to decode and inform how visual processing is controlled within the brain.

3. Decoding the brain's visual system

The challenge of creating artificial systems capable of emulating biological visual processing is formidable. However, recent efforts to understand and reverse engineer the brain's ventral visual stream, a series of interconnected cortical nuclei responsible for hierarchically processing and encoding of images into explicit neural representations, have shown great promise in the creation of robust AI systems capable of decoding and interpreting human visual processing, as well as performing complex visual intelligence skills including image recognition (Federer, Xu, Fyshe, & Zylberberg, 2020; Verschae & Ruiz-del-Solar, 2015), motion detection (Manchanda & Sharma, 2016; Wu, McGinnity, Maguire, Cai, & Valderrama-Gonzalez, 2008), and object tracking (Luo et al., 2020; Soleimanitaleb, Keyvanrad, & Jafari, 2019; Zhang et al., 2021).

In an effort to understand and measure human visual perception, machine learning models, including support-vector networks, have been trained to decode stimulus-induced fMRI activity patterns in the human V1 cortical area, and were able to visually reconstruct the local contrast of presented and internal mental images (Kamitani & Tong, 2005; Miyawaki et al., 2008). Similarly, those trained to decode stimulus-induced activity in higher visual cortical areas were able to identify the semantic contents of dream imagery (Horikawa, Tamaki, Miyawaki, & Kamitani, 2013). These findings indicate that the visual features of both perceived and mental images are represented in the same neural substrates (lower and higher visual areas for low-level perceptual and high-level semantic features, respectively), supporting previous evidence from human PET imaging studies revealing mental imagery to activate the primary visual cortex, an

area necessary for visual perception (Kosslyn et al., 1993; Kosslyn, Thompson, Klm, & Alpert, 1995). Additionally, these studies add to a growing literature demonstrating the utility of AI for decoding of brain imaging data for objective measurement of human visual experience (Kamitani & Tong, 2005; Nishimoto et al., 2011). Finally, beyond machine learning methods, the incorporation of a deep generator network (DGN) to a very deep convolutional neural network (CNN) image reconstruction method, allowing CNN hierarchical processing layers to be fully utilized in a manner similar to that of the human visual system, has recently been demonstrated to improve the quality of visual reconstructions of perceived or mental images compared with the same CNN without the DGN (Shen, Horikawa, Majima, & Kamitani, 2019).

Interestingly, neural networks trained to perform visual tasks have often been reported to acquire similar representations to regions of the brain's visual system required for the same tasks (Nonaka, Majima, Aoki, & Kamitani, 2020; Yamins & DiCarlo, 2016). CNNs incorporating hierarchical processing layers similar to that of the visual ventral stream and trained on image recognition tasks have been reported to be able to accurately predict neural responses in the inferior temporal (IT) cortex, the highest area of the ventral visual stream, of primates (Cadieu et al., 2014; Khaligh-Razavi & Kriegeskorte, 2014; Yamins et al., 2014). What is more, high-throughput computational evaluation of candidate CNN models revealed a strong correlation between a model's object recognition capability and its ability to predict IT cortex neural activity (Yamins et al., 2014). Accordingly, recent evidence revealed that the inclusion of components that closely predict the activity of the front-end of the visual stream (V1 area) improve the accuracy of CNNs by reducing their susceptibility to errors resulting from image perturbations, so-called white box adversarial attacks (Dapello et al., 2020).

While these studies appear to suggest the merit of "brain-like" AI systems for visual processing, until recently there has been no method to objectively measure how "brain-like" an AI visual processing model is. In response to this concern, two novel metrics, the Brain-Score (BS) (Schrimpf et al., 2020) and the brain hierarchy (BH) score (Nonaka et al., 2020), have been created to assess the functional similarity between AI models and the human visual system. Specifically, the BS measures the ability of models to predict brain activity and behavior, whereas the BH is designed to evaluate the hierarchical homology across layers/areas between neural networks and the brain (Nonaka et al., 2020; Schrimpf et al., 2020). Interestingly, while evaluation of several commonly used AI visual processing models found a positive correlation between the BS and the accuracy for image recognition (i.e., brain-like neural networks performed better), the opposite result was found when the BH was used (Nonaka et al., 2020; Schrimpf et al., 2020). Although these findings appear to contradict each other, more recently developed high-performance neural networks tended to have a lower BS, suggesting that AI vision may now be diverging from human vision (Schrimpf et al., 2020). Importantly, and particularly for the BS, it should be considered that while the ability of a model to predict brain activity may indicate its functional similarity, it does not necessarily mean that the model is emulating actual brain mechanisms. In fact, statisticians have long stressed this importance of the distinction between explanatory and predictive modeling (Shmueli, 2010). Thus, if we intend to use AI systems to model and possibly inform our understanding of visual processing in the brain, it is important that we continue to increase the structural and mechanistic correspondence of AI models to their counterpart systems within the brain, as well as strengthening the ability of metrics to measure such correspondence. Indeed, considering the known complexity of the brain's visual system, including the existence of multiple cell types (Gonchar, Wang,

& Burkhalter, 2008; Pfeffer, Xue, He, Huang, & Scanziani, 2013) that are modulated by various neurotransmitters (Azimi et al., 2020; Noudoost & Moore, 2011), it is likely that comparatively simplistic artificial neural networks do not yet come close to fully modeling the myriad of processes contributing to biological visual processing.

Finally, in addition to their utility for image recognition, brain-inspired neural networks are now beginning to be applied to innovative and practical uses within the field of visual neuroscience research. One example of this is the recent use of an artificial neural network to design precise visual patterns that can be projected directly onto the retina of primates to accurately control the activity of individual or groups of ventral stream (V4 area) neurons (Bashivan, Kar, & DiCarlo, 2019). These findings indicate the potential of this method for non-invasive control of neural activity in the visual cortex, creating a powerful tool for neuroscientists. In the next section, we further describe how AI is now increasingly being utilized for the advancement of neuroscience research, including in the objective analysis of animal behavior and its neural basis.

4. AI for analyzing behavior and its neural correlates

Understanding the relationship between neural activity and behavior is a critical goal of neuroscience. Recently developed large-scale neural imaging techniques have now enabled huge quantities of data to be collected during behavioral tasks in animals (Ahrens & Engert, 2015; Cardin, Crair, & Higley, 2020; Weisenburger & Vaziri, 2016; Yang & Yuste, 2017). However, given the quantity and speed of individual movements animals perform during behavioral tasks, as well as the difficulty in identifying individual neurons among large and crowded neural imaging datasets, it has been challenging for researchers to effectively and objectively analyze animal behavior and its precise neural correlates (Berman, 2018; Giovannucci et al., 2019; von Ziegler, Sturman, & Bohacek, 2021).

To address difficulties in human labeling of animal behavior, researchers have turned to AI for help. Over the last few years, several open-source, deep learning-based software toolboxes have been developed for 3D markerless pose estimation across several species and types of behaviors (Arac, Zhao, Dobkin, Carmichael, & Golshani, 2019; Forys, Xiao, Gupta, & Murphy, 2020; Graving et al., 2019; Günel et al., 2019; Mathis et al., 2018; Nath et al., 2019; Pereira et al., 2019). Perhaps the most widely used of these has been DeepLabCut, a deep neural network that incorporates the feature detectors from DeeperCut, a multi-person pose estimation model, and is able to accurately estimate the pose of several commonly used laboratory animals with minimal training (Lauer et al., 2021; Mathis et al., 2018; Nath et al., 2019). This pose estimation data can then be combined with various supervised machine learning tools, including JAABA (Kabra, Robie, Rivera-Alba, Branson, & Branson, 2013) and SimBA (Nilsson et al., 2020) that allow for automated identification of specific behaviors labeled by humans, such as grooming, freezing, and various social behaviors. This combined use of such tools has been shown to be able to match human ability for accurate quantification of several types of behaviors, and can outperform commercially available animal-tracking software packages (Sturman et al., 2020). In addition to supervised machine learning analysis of animal behavioral data, several unsupervised machine learning tools have been developed, including MotionMapper (Berman, Choi, Bialek, & Shaevitz, 2014), MoSeq (Wiltschko et al., 2015), and more recently, uBAM (Brattoli et al., 2021). These unsupervised approaches allow objective classification of the full repertoire of animal behavior and can potentially uncover subtle behavioral traits that might be missed by humans (Kwok, 2019).

As with animal behavioral data, human annotation of animal neural imaging datasets acquired from large-scale recording of neural activity, such as that acquired using in-vivo imaging of neural activity markers such as calcium indicators, is time consuming and suffers from a large degree of variability between annotators in the segmentation of individual neurons (Giovannucci et al., 2019). In the last decade, several tools utilizing classic machine learning and deep learning approaches have been developed to assist the analysis of animal calcium imaging data through the automated detection and quantification of individual neuron activity, known as source extraction (Pnevmatikakis, 2019). The most widely used of these have been unsupervised machine learning approaches employing activity-based segmentation algorithms, including principal component and independent component analysis (PCA/ICA) (Mukamel, Nimmerjahn, & Schnitzer, 2009), variations of constrained non-negative matrix factorization (CNMF) (Friedrich, Giovannucci, & Pnevmatikakis, 2021; Guan et al., 2018; Pnevmatikakis et al., 2016; Zhou et al., 2018), and dictionary learning (Giovannucci et al., 2017; Petersen, Simon, & Witten, 2017), to extract signals of neuron-like regions of interest from the background. While these techniques offer the benefit that they require no training and thus can be applied to analysis of various cell types and even dendritic imaging, they often suffer from false positives and are unable to identify low-activity neurons, making it difficult to longitudinally track the activity of neurons that may be temporally inactive in certain contexts (Lu et al., 2018). To address this limitation, several supervised deep learning approaches that segment neurons based upon features learned from human-labeled calcium imaging datasets have been developed (Apthorpe et al., 2016; Denis, Dard, Quiroli, Cossart, & Picardo, 2020; Giovannucci et al., 2019; Klibisz, Rose, Eicholtz, Blundon, & Zakharenko, 2017; Soltanian-Zadeh, Sahingur, Blau, Gong, & Farsiu, 2019; Xu, Su, Zhut, Guan, & Zhangt, 2016). Many of these tools, including U-Net2DS (Klibisz et al., 2017), STNeuroNet (Soltanian-Zadeh et al., 2019), and DeepCINAC (Denis et al., 2020), train a CNN to segment neurons in either 2D or 3D space and have been demonstrated to be able to detect neurons with near-human accuracy, and outperform other techniques including PCA/ICA, allowing for accurate, fast, and reproducible neural detection and classification (Apthorpe et al., 2016; Giovannucci et al., 2019; Mukamel et al., 2009).

Finally, efforts are now being made to combine AI analysis of animal behavior and neural imaging data, not only for automated mapping of behavior to its neural correlates, but in order to predict and model animal behavior based on analyzed neural activity data. One such recently developed system is BehaveNet, a probabilistic framework for the unsupervised analysis of behavioral video, with semi-supervised decoding of neural activity (Batty et al., 2019). The resulting generative models of this framework are able to decode animal neural activity data and create probabilistic full-resolution video simulations of behavior (Batty et al., 2019). Further development of technologies designed to automate the mapping of neural activity patterns with behavioral motifs may help to elucidate how discrete patterns of neural activity are related to specific movements (Musall et al., 2019).

While the studies here describe approaches to analyzing and modeling healthy behavior and brain activity in animals, efforts have also been made to use AI to understand and identify abnormal brain functioning. In the next section, we discuss AI-based approaches to objective classification of psychiatric disorders, and how deep learning approaches have been utilized for modeling such disorders in artificial neural networks.

5. The interface between AI and psychiatry

Despite the adoption of standardized diagnostic criteria in clinical manuals such as the Diagnostic and Statistical Manual of Mental Disorders (DSM) and the International Classification of Disease (ICD), psychiatric and developmental disorders are still primarily identified based upon a patient's subjective behavioral symptoms and self-report measures. Not only is this method often unreliable due to its subjectivity (Wakefield, 2016), but it also leads to an explanatory gap between phenomenology and neurobiology. However, in the last few decades, huge advancements in the power of computing, alongside the collection of large neuroimaging datasets, have allowed researchers to begin to bridge this gap by using AI to identify, model, and potentially even treat psychiatric and developmental disorders.

One area of particular promise has been the use of AI for the objective identification of brain disorders. Using machine learning methods, classifiers have been built to predict diagnostic labels of psychiatric and developmental disorders (see Bzdok & Meyer-Lindenberg, 2017; Cho et al., 2019; Zhou et al., 2020 for review). The scores produced by these probabilistic classifiers provide a degree of classification certainty that can be interpreted as a neural liability for the disorder and represent new biological dimensions of the disorders. However, while many of these classifiers, including those for schizophrenia (Greenstein, Malley, Weisinger, Clasen, & Gogtay, 2012; Orrù, Pettersson-Yeo, Marquand, Sartori, & Mechelli, 2012; Yassin et al., 2020) and ASD (Eslami, Almuqhim, Raiker, & Saeed, 2021; Yassin et al., 2020), are able to accurately identify the intended disorder, a major criticism has been that they are often only validated in a single sample cohort. To address this issue, recent attempts have been made to establish robust classifiers using larger and more varied sample data. This has led to the identification of classifiers for ASD and schizophrenia that could be generalized to independent cohorts, regardless of ethnicity, country, and MRI vendor, and still demonstrated classification accuracies of between 61%–76% (Yamada et al., 2017; Yoshihara et al., 2020). Asides from machine learning, deep learning approaches have also been applied to the classification of psychiatric and developmental disorders (see Durstewitz et al., 2019; Koppe et al., 2021, for review). A major advantage of deep neural networks is that their multi-layered design makes them particularly suited to learning high-level representations from complex raw data, allowing features to extracted from neuroimaging data with far less parameters than machine learning architectures (Durstewitz et al., 2019; Jang, Plis, Calhoun, & Lee, 2017; Koppe et al., 2021; Plis et al., 2014; Schmidhuber, 2015). Accordingly, in the last few years, several deep neural networks have been reported to effectively classify brain disorders from neuroimaging data, including schizophrenia (Oh, Oh, Lee, Chae, & Yun, 2020; Sun et al., 2021; Yan et al., 2019; Zeng et al., 2018), autism (Guo et al., 2017; Heinsfeld, Franco, Craddock, Buchweitz, & Meneguzzi, 2018; Misman et al., 2019; Raj & Masood, 2020), ADHD (Chen, Li, et al., 2019; Chen, Song, & Li, 2019; Dubreuil-Vall, Ruffini, & Camprodon, 2020), and depression (Li, La, Wang, Hu, & Zhang, 2020; Uyulan et al., 2020). Further development of AI models for data-driven, dimensional psychiatry will likely help to address the current discontent surrounding categorical diagnostic criteria.

In addition to classification of disorders based on neuroimaging data, AI is also increasingly being used to model various psychiatric and developmental disorders (see Lanillos et al., 2020 for in-depth reviews). This largely began in the 1980s and 90s with studies modeling schizophrenia and ASD using artificial neural networks (Cohen, 1994; Cohen & Servan-Schreiber, 1992; Hoffman, 1987; Horn & Ruppin, 1995). Many of these models were inspired by biological evidence of structural and synaptic abnormalities associated with particular psychiatric disorder

symptoms. For example, evidence of reduced metabolism in the frontal cortex (Feinberg, 1983; Feinberg, Koresko, & Gottlieb, 1965; Feinberg, Koresko, Gottlieb, & Wender, 1964) and aberrant synaptic regeneration of abnormal brain structures (Stevens, 1992) have prompted the creation of neural networks designed to simulate how synaptic pruning (Hoffman & Dobscha, 1989) and reactive synaptic reorganization (Horn & Ruppin, 1995; Ruppin, Reggia, & Horn, 1996) may explain delusions and hallucinations in schizophrenia patients. Similarly, neural network models of excessive or reduced neuronal connections (Cohen, 1994, 1998; Thomas, Knowland, & Karmiloff-Smith, 2011) were generated to model biological observations of abnormal neuronal density in cortical, limbic and cerebellar regions (Bailey et al., 1998; Bauman, 1991; Bauman & Kemper, 1985) hypothesized to contribute to developmental regression in ASD. Excitingly, more recently, deep learning models, including high-dimensional RNN models of schizophrenia (Yamashita & Tani, 2012) and ASD (Idei et al., 2017, 2018), have begun to be implemented into robots, allowing direct observation and comparison of modeled behavior with that seen in patients.

Finally, in the near future, AI could begin to play an important role in the treatment of psychiatric and developmental disorders. Computer-assisted therapy (CAT), including AI chatbots delivering cognitive behavioral therapies, is beginning to be tested for treatment of psychiatric disorders including depression and anxiety (Carroll & Rounsaville, 2010; Fitzpatrick, Darcy, & Vierhile, 2017; Fulmer, Joerin, Gentile, Lakerink, & Rauws, 2018). While still in their infancy, these CATs offer distinct advantages to human-led therapies in terms of price and accessibility, although their effectiveness in comparison to currently used therapeutic methods has yet to be robustly measured. Additionally, the identification of neuroimaging-based classifiers for psychiatric disorders (described above) has inspired the launch of a real-time fMRI-based neurofeedback projects in which patients attempt to normalize their own brain connectivity pattern through neurofeedback. Meta-analyses of such studies have indicated neurofeedback treatments to result in significant amelioration of the symptoms of several disorders including schizophrenia, depression, anxiety disorder, and ASD, suggesting the potential benefit of further use of such digital treatments (Dudek & Dodell-Feder, 2020; Schoenberg & David, 2014).

6. Conclusions

Since the inception of AI research midway through the last century, the brain has served as the primary source of inspiration for the creation of artificial systems of intelligence. This is largely based upon the reasoning that the brain is proof of concept of a comprehensive intelligence system capable of perception, planning, and decision making, and therefore offers an attractive template for the design of AI. In this review, based upon topics presented at the 2020 International Symposium on Artificial Intelligence and Brain Science, we have discussed how brain-inspired mechanistic, structural, and functional elements are being utilized to create novel, and optimize existing, AI systems. In particular, this has led to the development of high-dimensional deep neural networks often incorporating hierarchical architectures inspired by those found in the brain and capable of feats of intelligence including visual object recognition and memory-based cognitive tasks. Advancements in AI have also helped to foster progress within the field of neuroscience. Here we have described how the use of machine learning and neural networks for automated analysis of big data has revolutionized the analysis of animal behavioral and neuroimaging studies, as well as being utilized for objective classification of psychiatric and developmental disorders.

Importantly, while it has not been discussed in great detail in the current review, it should be considered that the relationship between AI and neuroscience is not simply two-way, but rather also includes the field of cognitive science (see Battleday, Peterson, & Griffiths, 2021; Cichy & Kaiser, 2019; Forbus, 2010; Kriegeskorte & Douglas, 2018, for review). Indeed, over the years, much of AI research has been guided by theories of brain functioning established by cognitive scientists (Elman, 1990; Hebb, 1949; Rumelhart & McClelland, 1986). For example, the convolutional neural networks discussed earlier in this review (in the section on visual processing) were inspired in part by computational models of cognition within the brain, including principles such as nonlinear feature maps and pooling of inputs, which were themselves derived from observations from neurophysiological studies in animals (Battleday et al., 2021; Fukushima, 1980; Hubel & Wiesel, 1962; Mozer, 1987; Riesenhuber & Poggio, 1999). In turn, neural networks have been used to guide new cognitive models of intellectual abilities, including perception, memory, and language, giving rise to the connectionism movement within cognitive science (Barrow, 1996; Fodor & Pylyshyn, 1988; Mayor, Gomez, Chang, & Lupyan, 2014; Yamins & DiCarlo, 2016). If we are to use AI to model and potentially elucidate brain functioning, the primary focus of cognitive science, it is important that we continue to not only use biological data from neuroscience studies, but also cognitive models, to inspire the architectural, mechanistic, and algorithmic design of artificial neural networks.

Despite their accomplishments and apparent complexity, current AI systems are still remarkably simplistic in comparison to brain networks and in many cases still lack the ability to accurately model brain functions (Bae, Kim, & Kim, 2021; Barrett, Morcos, & Macke, 2019; Hasson et al., 2020; Pulvermüller, Tomasello, Henningsen-Schomers, & Wennekers, 2021; Tang et al., 2019). A major limitation is that, in general, current models are still not able to model the brain at multiple levels; from synaptic reorganization and the influence of neurotransmitter and hormone neuromodulation of neuronal excitability at the microlevel, to large-scale synchronization of spiking activity and global connectivity at the macrolevel. In fact, integration of various AI models of brain functioning, including the models of the cerebral cortex described in this review, as well as models of other brain regions, including limbic and motor control regions (Kowalczuk & Czubenko, 2016; Merel, Botvinick, & Wayne, 2019; Parsapoor, 2016), remains one of the greatest challenges in the creation of an AGI system capable of modeling the entire brain. In spite of these difficulties, it is clear that the continued interaction between neuroscience and AI will undoubtedly expedite progress in both areas.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

References

Ahrens, M. B., & Engert, F. (2015). Large-scale imaging in small brains. *Current Opinion in Neurobiology*, *32*, 78–86. http://dx.doi.org/10.1016/j.conb.2015.01.007.

Amit, D. J., Bernacchia, A., & Yakovlev, V. (2003). Multiple-object working memory—A model for behavioral performance. *Cerebral Cortex*, *13*(5), 435–443. http://dx.doi.org/10.1093/cercor/13.5.435.

Amit, D. J., & Brunel, N. (1995). Learning internal representations in an attractor neural network with analogue neurons. *Network. Computation in Neural Systems*, *6*(3), 359–388. http://dx.doi.org/10.1088/0954-898x_6_3_004.

Apthorpe, N. J., Riordan, A. J., Aguilar, R. E., Homann, J., Gu, Y., Tank, D. W., et al. (2016). Automatic neuron detection in calcium imaging data using convolutional networks. ArXiv.

Arac, A., Zhao, P., Dobkin, B. H., Carmichael, S. T., & Golshani, P. (2019). DeepBehavior: A deep learning toolbox for automated analysis of animal and human behavior imaging data. *Frontiers in Systems Neuroscience*, *13*, 20. http://dx.doi.org/10.3389/fnsys.2019.00020.

Azimi, Z., Barzan, R., Spoida, K., Surdin, T., Wollenweber, P., Mark, M. D., et al. (2020). Separable gain control of ongoing and evoked activity in the visual cortex by serotonergic input. *ELife*, *9*, Article e53552. http://dx.doi.org/10.7554/elife.53552.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829–839. http://dx.doi.org/10.1038/nrn1201.

Bae, H., Kim, S. J., & Kim, C.-E. (2021). Lessons from deep neural networks for studying the coding principles of biological neural networks. *Frontiers in Systems Neuroscience*, *14*, Article 615129. http://dx.doi.org/10.3389/fnsys.2020.615129.

Bailey, A., Luthert, P., Dean, A., Harding, B., Janota, I., Montgomery, M., et al. (1998). A clinicopathological study of autism. *Brain*, *121*(5), 889–905. http://dx.doi.org/10.1093/brain/121.5.889.

Barrett, D. G., Morcos, A. S., & Macke, J. H. (2019). Analyzing biological and artificial neural networks: Challenges with opportunities for synergy? *Current Opinion in Neurobiology*, *55*, 55–64. http://dx.doi.org/10.1016/j.conb.2019.01.007.

Barrow, H. (1996). Connectionism and neural networks. In M. A. Boden (Ed.), *Artificial intelligence: Handbook of perception and cognition*. Academic Press, http://dx.doi.org/10.1016/b978-012161964-0/50007-8.

Bashivan, P., Kar, K., & DiCarlo, J. J. (2019). Neural population control via deep image synthesis. *Science*, *364*(6439), Article eaav9436. http://dx.doi.org/10.1126/science.aav9436.

Bastos, A. M., Loonis, R., Kornblith, S., Lundqvist, M., & Miller, E. K. (2018). Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proceedings of the National Academy of Sciences*, *115*(5), 1117–1122. http://dx.doi.org/10.1073/pnas.1710323115.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695–711. http://dx.doi.org/10.1016/j.neuron.2012.10.038.

Battleday, R. M., Peterson, J. C., & Griffiths, T. L. (2021). From convolutional neural networks to models of higher-level cognition (and back again). *Annals of the New York Academy of Sciences*, http://dx.doi.org/10.1111/nyas.14593.

Batty, E., Whiteway, M., Saxena, S., Biderman, D., Abe, T., Musall, S., et al. (2019). BehaveNet: Nonlinear embedding and Bayesian neural decoding of behavioral videos. In *33rd conference on neural information processing systems*.

Bauman, M. L. (1991). Microscopic neuroanatomic abnormalities in autism. *Pediatrics*, *87*(5 Pt 2), 791–796.

Bauman, M., & Kemper, T. L. (1985). Histoanatomic observations of the brain in early infantile autism. *Neurology*, *35*(6), 866. http://dx.doi.org/10.1212/wnl.35.6.866.

Beiser, D. G., & Houk, J. C. (1998). Model of cortical-basal ganglionic processing: Encoding the serial order of sensory events. *Journal of Neurophysiology*, *79*(6), 3168–3188. http://dx.doi.org/10.1152/jn.1998.79.6.3168.

Bengio, Y. (1993). A connectionist approach to speech recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, *07*(04), 647–667. http://dx.doi.org/10.1142/s0218001493000327.

Berman, G. J. (2018). Measuring behavior across scales. *BMC Biology*, *16*(1), 23. http://dx.doi.org/10.1186/s12915-018-0494-7.

Berman, G. J., Choi, D. M., Bialek, W., & Shaevitz, J. W. (2014). Mapping the stereotyped behaviour of freely moving fruit flies. *Journal of the Royal Society Interface*, *11*(99), Article 20140672. http://dx.doi.org/10.1098/rsif.2014.0672.

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: Insights from spatial processing. *Nature Reviews Neuroscience*, *9*(3), 182–194. http://dx.doi.org/10.1038/nrn2335.

Blumenfeld, R. S., & Ranganath, C. (2007). Prefrontal cortex and long-term memory encoding: An integrative review of findings from neuropsychology and neuroimaging. *The Neuroscientist*, *13*(3), 280–291. http://dx.doi.org/10.1177/1073858407299290.

Boran, E., Fedele, T., Klaver, P., Hilfiker, P., Stieglitz, L., Grunwald, T., et al. (2019). Persistent hippocampal neural firing and hippocampal-cortical coupling predict verbal working memory load. *Science Advances*, *5*(3), Article eaav3687. http://dx.doi.org/10.1126/sciadv.aav3687.

Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review*, *113*(2), 201–233. http://dx.doi.org/10.1037/0033-295x.113.2.201.

Brattoli, B., Büchler, U., Dorkenwald, M., Reiser, P., Filli, L., Helmchen, F., et al. (2021). Unsupervised behaviour analysis and magnification (uBAM) using deep learning. *Nature Machine Intelligence*, *3*(6), 495–506. http://dx.doi.org/10.1038/s42256-021-00326-x.

Bzdok, D., & Meyer-Lindenberg, A. (2017). Machine learning for precision psychiatry: Opportunities and challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(3), 223–230. http://dx.doi.org/10.1016/j.bpsc.2017.11.007.

Cadieu, C. F., Hong, H., Yamins, D. L. K., Pinto, N., Ardila, D., Solomon, E. A., et al. (2014). Deep neural networks rival the representation of primate IT cortex for core visual object recognition. *PLoS Computational Biology*, *10*(12), Article e1003963. http://dx.doi.org/10.1371/journal.pcbi.1003963.

Cardin, J. A., Crair, M. C., & Higley, M. J. (2020). Mesoscopic imaging: Shining a wide light on large-scale neural dynamics. *Neuron*, *108*(1), 33–43. http://dx.doi.org/10.1016/j.neuron.2020.09.031.

Carroll, K. M., & Rounsaville, B. J. (2010). Computer-assisted therapy in psychiatry: Be brave—It's a new world. *Current Psychiatry Reports*, *12*(5), 426–432. http://dx.doi.org/10.1007/s11920-010-0146-2.

Chaudhuri, R., Knoblauch, K., Gariel, M.-A., Kennedy, H., & Wang, X.-J. (2015). A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron*, *88*(2), 419–431. http://dx.doi.org/10.1016/j.neuron.2015.09.008.

Chen, M., Li, H., Wang, J., Dillman, J. R., Parikh, N. A., & He, L. (2019). A multichannel deep neural network model analyzing multiscale functional brain connectome data for attention deficit hyperactivity disorder detection. *Radiology: Artificial Intelligence*, *2*(1), Article e190012. http://dx.doi.org/10.1148/ryai.2019190012.

Chen, H., Song, Y., & Li, X. (2019). A deep learning framework for identifying children with ADHD using an EEG-based brain network. *Neurocomputing*, *356*, 83–96. http://dx.doi.org/10.1016/j.neucom.2019.04.058.

Cho, G., Yim, J., Choi, Y., Ko, J., & Lee, S.-H. (2019). Review of machine learning algorithms for diagnosing mental illness. *Psychiatry Investigation*, *16*(4), 262–269. http://dx.doi.org/10.30773/pi.2018.12.21.2.

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J.-D. (2017). The distributed nature of working memory. *Trends in Cognitive Sciences*, *21*(2), 111–124. http://dx.doi.org/10.1016/j.tics.2016.12.007.

Cichy, R. M., & Kaiser, D. (2019). Deep neural networks as scientific models. *Trends in Cognitive Sciences*, *23*(4), 305–317. http://dx.doi.org/10.1016/j.tics.2019.01.009.

Cohen, I. L. (1994). An artificial neural network analogue of learning in autism. *Biological Psychiatry*, *36*(1), 5–20. http://dx.doi.org/10.1016/0006-3223(94)90057-4.

Cohen, I. L. (1998). In D. J. Stein, & J. Ludik (Eds.), *Neural network analysis of learning in autism*. Cambridge University Press.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review*, *99*(1), 45–77. http://dx.doi.org/10.1037/0033-295x.99.1.45.

Crick, F. (1989). The recent excitement about neural networks. *Nature*, *337*(6203), 129–132. http://dx.doi.org/10.1038/337129a0.

Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C. K., Hassabis, D., Munos, R., et al. (2020). A distributional code for value in dopamine-based reinforcement learning. *Nature*, *577*(7792), 671–675. http://dx.doi.org/10.1038/s41586-019-1924-6.

Dapello, J., Marques, T., Schrimpf, M., Geiger, F., Cox, D. D., & DiCarlo, J. J. (2020). Simulating a primary visual cortex at the front of CNNs improves robustness to image perturbations. http://dx.doi.org/10.1101/2020.06.16.154542, BioRxiv, 2020.06.16.154542.

Denis, J., Dard, R. F., Quiroli, E., Cossart, R., & Picardo, M. A. (2020). DeepCINAC: A deep-learning-based Python toolbox for inferring calcium imaging neuronal activity based on movie visualization. *ENeuro*, *7*(4), http://dx.doi.org/10.1523/eneuro.0038-20.2020, ENEURO.0038-20.2020.

Diehl, P. U., Zarrella, G., Cassidy, A., Pedroni, B. U., & Neftci, E. (2016). Conversion of artificial recurrent neural networks to spiking neural networks for low-power neuromorphic hardware. ArXiv.

Diesmann, M., Gewaltig, M.-O., & Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature*, *402*(6761), 529–533. http://dx.doi.org/10.1038/990101.

Dubreuil-Vall, L., Ruffini, G., & Camprodon, J. A. (2020). Deep learning convolutional neural networks discriminate adult ADHD from healthy individuals on the basis of event-related spectral EEG. *Frontiers in Neuroscience*, *14*, 251. http://dx.doi.org/10.3389/fnins.2020.00251.

Dudek, E., & Dodell-Feder, D. (2020). The efficacy of real-time functional magnetic resonance imaging neurofeedback for psychiatric illness: A meta-analysis of brain and behavioral outcomes. *Neuroscience & Biobehavioral Reviews*, *121*, 291–306. http://dx.doi.org/10.1016/j.neubiorev.2020.12.020.

Durstewitz, D., Koppe, G., & Meyer-Lindenberg, A. (2019). Deep neural networks in psychiatry. *Molecular Psychiatry*, *24*(11), 1583–1598. http://dx.doi.org/10.1038/s41380-019-0365-9.

Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000). Neurocomputational models of working memory. *Nature Neuroscience*, *3*(Suppl. 11), 1184–1191. http://dx.doi.org/10.1038/81460.

Eichenbaum, H. (2017). Prefrontal–hippocampal interactions in episodic memory. *Nature Reviews Neuroscience*, 18(9), 547–558. http://dx.doi.org/10.1038/nrn.2017.74.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179–211. http://dx.doi.org/10.1016/0364-0213(90)90002-e.

Eslami, T., Almuqhim, F., Raiker, J. S., & Saeed, F. (2021). Machine learning methods for diagnosing autism spectrum disorder and attention-deficit/hyperactivity disorder using functional and structural MRI: A survey. *Frontiers in Neuroinformatics*, 14, Article 575999. http://dx.doi.org/10.3389/fninf.2020.575999.

Federer, C., Xu, H., Fyshe, A., & Zylberberg, J. (2020). Improved object recognition using neural networks trained to mimic the brain's statistical properties. *Neural Networks*, 131, 103–114. http://dx.doi.org/10.1016/j.neunet.2020.07.013.

Feinberg, I. (1983). Schizophrenia: Caused by a fault in programmed synaptic elimination during adolescence? *Journal of Psychiatric Research*, 17(4), 319–334. http://dx.doi.org/10.1016/0022-3956(82)90038-3.

Feinberg, I., Koresko, R. L., & Gottlieb, F. (1965). Further observations on electrophysiological sleep patterns in schizophrenia. *Comprehensive Psychiatry*, 6(1), 21–24. http://dx.doi.org/10.1016/s0010-440x(65)80004-9.

Feinberg, I., Koresko, R. L., Gottlieb, F., & Wender, P. H. (1964). Sleep electroencephalographic and eye-movement patterns in schizophrenic patients. *Comprehensive Psychiatry*, 5(1), 44–53. http://dx.doi.org/10.1016/s0010-440x(64)80042-0.

Fiebig, F., & Lansner, A. (2017). A spiking working memory model based on hebbian short-term potentiation. *Journal of Neuroscience*, 37(1), 83–96. http://dx.doi.org/10.1523/jneurosci.1989-16.2016.

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): A randomized controlled trial. *JMIR Mental Health*, 4(2), Article e19. http://dx.doi.org/10.2196/mental.7785.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1–2), 3–71. http://dx.doi.org/10.1016/0010-0277(88)90031-5.

Forbus, K. D. (2010). AI and cognitive science: The past and next 30 years. *Topics in Cognitive Science*, 2(3), 345–356. http://dx.doi.org/10.1111/j.1756-8765.2010.01083.x.

Forys, B. J., Xiao, D., Gupta, P., & Murphy, T. H. (2020). Real-time selective markerless tracking of forepaws of head fixed mice using deep neural networks. *ENeuro*, 7(3), http://dx.doi.org/10.1523/eneuro.0096-20.2020, ENEURO.0096-20.2020.

Friedrich, J., Giovannucci, A., & Pnevmatikakis, E. A. (2021). Online analysis of microendoscopic 1-photon calcium imaging data streams. *PLoS Computational Biology*, 17(1), Article e1008565. http://dx.doi.org/10.1371/journal.pcbi.1008565.

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4), 193–202. http://dx.doi.org/10.1007/bf00344251.

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using psychological artificial intelligence (tess) to relieve symptoms of depression and anxiety: Randomized controlled trial. *JMIR Mental Health*, 5(4), Article e64. http://dx.doi.org/10.2196/mental.9782.

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, 173(3997), 652–654. http://dx.doi.org/10.1126/science.173.3997.652.

Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B. L., Koay, S. A., et al. (2019). CaImAn an open source tool for scalable calcium imaging data analysis. *ELife*, 8, Article e38173. http://dx.doi.org/10.7554/elife.38173.

Giovannucci, A., Friedrich, J., Kaufman, M., Churchland, A., Chklovskii, D., Paninski, L., et al. (2017). OnACID: Online analysis of calcium imaging data in real time*. http://dx.doi.org/10.1101/193383, BioRxiv, 193383.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, 14(3), 477–485. http://dx.doi.org/10.1016/0896-6273(95)90304-6.

Gonchar, Y., Wang, Q., & Burkhalter, A. H. (2008). Multiple distinct subtypes of GABAergic neurons in mouse visual cortex identified by triple immunostaining. *Frontiers in Neuroanatomy*, 2, 3. http://dx.doi.org/10.3389/neuro.05.003.2007.

Graving, J. M., Chae, D., Naik, H., Li, L., Koger, B., Costelloe, B. R., et al. (2019). DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *ELife*, 8, Article e47994. http://dx.doi.org/10.7554/elife.47994.

Greenstein, D., Malley, J. D., Weisinger, B., Clasen, L., & Gogtay, N. (2012). Using multivariate machine learning methods and structural MRI to classify childhood onset schizophrenia and healthy controls. *Frontiers in Psychiatry*, 3, 53. http://dx.doi.org/10.3389/fpsyt.2012.00053.

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science*, 11(1), 23–63. http://dx.doi.org/10.1111/j.1551-6708.1987.tb00862.x.

Guan, J., Li, J., Liang, S., Li, R., Li, X., Shi, X., et al. (2018). NeuroSeg: Automated cell detection and segmentation for in vivo two-photon Ca2+ imaging data. *Brain Structure and Function*, 223(1), 519–533. http://dx.doi.org/10.1007/s00429-017-1545-5.

Günel, S., Rhodin, H., Morales, D., Campagnolo, J., Ramdya, P., & Fua, P. (2019). DeepFly3D, a deep learning-based approach for 3D limb and appendage tracking in tethered, adult Drosophila. *ELife*, 8, Article e48571. http://dx.doi.org/10.7554/elife.48571.

Guo, X., Dominick, K. C., Minai, A. A., Li, H., Erickson, C. A., & Lu, L. J. (2017). Diagnosing autism spectrum disorder from brain resting-state functional connectivity patterns using a deep neural network with a novel feature selection method. *Frontiers in Neuroscience*, 11, 460. http://dx.doi.org/10.3389/fnins.2017.00460.

Han, B., Sengupta, A., & Roy, K. (2016). On the energy benefits of spiking deep neural networks: A case study (special session paper). In *2016 international joint conference on neural networks* (pp. 971–976). http://dx.doi.org/10.1109/ijcnn.2016.7727303.

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron*, 95(2), 245–258. http://dx.doi.org/10.1016/j.neuron.2017.06.011.

Hasson, U., Nastase, S. A., & Goldstein, A. (2020). Direct fit to nature: An evolutionary perspective on biological and artificial neural networks. *Neuron*, 105(3), 416–434. http://dx.doi.org/10.1016/j.neuron.2019.12.002.

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2007). Towards an executive without a homunculus: Computational models of the prefrontal cortex/basal ganglia system. *Philosophical Transactions of the Royal Society, Series B (Biological Sciences)*, 362(1485), 1601–1613. http://dx.doi.org/10.1098/rstb.2007.2055.

Hebb, D. O. (1949). *The organization of behavior*. Wiley & Sons.

Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., & Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage: Clinical*, 17, 16–23. http://dx.doi.org/10.1016/j.nicl.2017.08.017.

Hoffman, R. E. (1987). Computer simulations of neural information processing and the schizophrenia-mania dichotomy. *Archives of General Psychiatry*, 44(2), 178–188. http://dx.doi.org/10.1001/archpsyc.1987.01800140090014.

Hoffman, R. E., & Dobscha, S. K. (1989). Cortical pruning and the development of schizophrenia: A computer model. *Schizophrenia Bulletin*, 15(3), 477–490. http://dx.doi.org/10.1093/schbul/15.3.477.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8), 2554–2558. http://dx.doi.org/10.1073/pnas.79.8.2554.

Horikawa, T., Tamaki, M., Miyawaki, Y., & Kamitani, Y. (2013). Neural decoding of visual imagery during sleep. *Science*, 340(6132), 639–642. http://dx.doi.org/10.1126/science.1234330.

Horn, D., & Ruppin, E. (1995). Compensatory mechanisms in an attractor neural network model of schizophrenia. *Neural Computation*, 7(1), 182–205. http://dx.doi.org/10.1162/neco.1995.7.1.182.

Huang, Q.-S., & Wei, H. (2021). A computational model of working memory based on spike-timing-dependent plasticity. *Frontiers in Computational Neuroscience*, 15, Article 630999. http://dx.doi.org/10.3389/fncom.2021.630999.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal Physiology*, 160(1), 106–154. http://dx.doi.org/10.1113/jphysiol.1962.sp006837.

Idei, H., Murata, S., Chen, Y., Yamashita, Y., Tani, J., & Ogata, T. (2017). Reduced behavioral flexibility by aberrant sensory precision in autism spectrum disorder: A neurorobotics experiment. In *2017 joint IEEE international conference on development and learning and epigenetic robotics* (pp. 271–276). http://dx.doi.org/10.1109/devlrn.2017.8329817.

Idei, H., Murata, S., Chen, Y., Yamashita, Y., Tani, J., & Ogata, T. (2018). A neurorobotics simulation of autistic behavior induced by unusual sensory precision. *Computational Psychiatry*, 2(0), 164–182. http://dx.doi.org/10.1162/cpsy_a_00019.

Jang, H., Plis, S. M., Calhoun, V. D., & Lee, J.-H. (2017). Task-specific feature extraction and classification of fMRI volumes using a deep neural network initialized with a deep belief network: Evaluation using sensorimotor tasks. *NeuroImage*, 145(Pt B), 314–328. http://dx.doi.org/10.1016/j.neuroimage.2016.04.003.

Kabra, M., Robie, A. A., Rivera-Alba, M., Branson, S., & Branson, K. (2013). JAABA: Interactive machine learning for automatic annotation of animal behavior. *Nature Methods*, 10(1), 64–67. http://dx.doi.org/10.1038/nmeth.2281.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8(5), 679–685. http://dx.doi.org/10.1038/nn1444.

Khaligh-Razavi, S.-M., & Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLoS Computational Biology*, 10(11), Article e1003915. http://dx.doi.org/10.1371/journal.pcbi.1003915.

Kim, R., Li, Y., & Sejnowski, T. J. (2019). Simple framework for constructing functional spiking recurrent neural networks. *Proceedings of the National Academy of Sciences*, 116(45), 22811–22820. http://dx.doi.org/10.1073/pnas.1905926116.

Kim, R., & Sejnowski, T. J. (2021). Strong inhibitory signaling underlies stable temporal dynamics and working memory in spiking neural networks. *Nature Neuroscience*, *24*(1), 129–139. http://dx.doi.org/10.1038/s41593-020-00753-w.

Klibisz, A., Rose, D., Eicholtz, M., Blundon, J., & Zakharenko, S. (2017). Fast, simple calcium imaging segmentation with fully convolutional networks. ArXiv.

Koppe, G., Meyer-Lindenberg, A., & Durstewitz, D. (2021). Deep learning for small and big data in psychiatry. *Neuropsychopharmacology*, *46*(1), 176–190. http://dx.doi.org/10.1038/s41386-020-0767-z.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Maljkovic, V., Weise, S. B., Chabris, C. F., et al. (1993). Visual mental imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience*, *5*(3), 263–287. http://dx.doi.org/10.1162/jocn.1993.5.3.263.

Kosslyn, S. M., Thompson, W. L., Klm, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, *378*(6556), 496–498. http://dx.doi.org/10.1038/378496a0.

Kowalczuk, Z., & Czubenko, M. (2016). Computational approaches to modeling artificial emotion – An overview of the proposed solutions. *Frontiers in Robotics and AI*, *3*, 21. http://dx.doi.org/10.3389/frobt.2016.00021.

Kriegeskorte, N., & Douglas, P. K. (2018). Cognitive computational neuroscience. *Nature Neuroscience*, *21*(9), 1148–1160. http://dx.doi.org/10.1038/s41593-018-0210-5.

Krotov, D., & Hopfield, J. J. (2019). Unsupervised learning by competing hidden units. *Proceedings of the National Academy of Sciences*, *116*(16), Article 201820458. http://dx.doi.org/10.1073/pnas.1820458116.

Kwok, R. (2019). Deep learning powers a motion-tracking revolution. *Nature*, *574*(7776), 137–138. http://dx.doi.org/10.1038/d41586-019-02942-5.

Lanillos, P., Oliva, D., Philippsen, A., Yamashita, Y., Nagai, Y., & Cheng, G. (2020). A review on neural network models of schizophrenia and autism spectrum disorder. *Neural Networks*, *122*, 338–363. http://dx.doi.org/10.1016/j.neunet.2019.10.014.

Lauer, J., Zhou, M., Ye, S., Menegas, W., Nath, T., Rahman, M. M., et al. (2021). Multi-animal pose estimation and tracking with DeepLabCut. http://dx.doi.org/10.1101/2021.04.30.442096, BioRxiv, 2021.04.30.442096.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. http://dx.doi.org/10.1038/nature14539.

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., et al. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, *1*(4), 541–551. http://dx.doi.org/10.1162/neco.1989.1.4.541.

Li, X., La, R., Wang, Y., Hu, B., & Zhang, X. (2020). A deep learning approach for mild depression recognition based on functional connectivity using electroencephalography. *Frontiers in Neuroscience*, *14*, 192. http://dx.doi.org/10.3389/fnins.2020.00192.

Lillicrap, T. P., Cownden, D., Tweed, D. B., & Akerman, C. J. (2016). Random synaptic feedback weights support error backpropagation for deep learning. *Nature Communications*, *7*(1), 13276. http://dx.doi.org/10.1038/ncomms13276.

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, *21*(6), 335–346. http://dx.doi.org/10.1038/s41583-020-0277-3.

Lisman, J. E., Fellous, J.-M., & Wang, X.-J. (1998). A role for NMDA-receptor channels in working memory. *Nature Neuroscience*, *1*(4), 273–275. http://dx.doi.org/10.1038/1086.

Lu, J., Li, C., Singh-Alvarado, J., Zhou, Z. C., Fröhlich, F., Mooney, R., et al. (2018). MIN1PIPE: A miniscope 1-photon-based calcium imaging signal extraction pipeline. *Cell Reports*, *23*(12), 3673–3684. http://dx.doi.org/10.1016/j.celrep.2018.05.062.

Lundqvist, M., Herman, P., & Miller, E. K. (2018). Working memory: Delay activity, yes! persistent activity? Maybe not. *Journal of Neuroscience*, *38*(32), 7013–7019. http://dx.doi.org/10.1523/jneurosci.2485-17.2018.

Lundqvist, M., Herman, P., Warden, M. R., Brincat, S. L., & Miller, E. K. (2018). Gamma and beta bursts during working memory readout suggest roles in its volitional control. *Nature Communications*, *9*(1), 394. http://dx.doi.org/10.1038/s41467-017-02791-8.

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., & Miller, E. K. (2016). Gamma and beta bursts underlie working memory. *Neuron*, *90*(1), 152–164. http://dx.doi.org/10.1016/j.neuron.2016.02.028.

Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W., & Kim, T.-K. (2020). Multiple object tracking: A literature review. *Artificial Intelligence*, *293*, Article 103448. http://dx.doi.org/10.1016/j.artint.2020.103448.

Manchanda, S., & Sharma, S. (2016). Analysis of computer vision based techniques for motion detection. In *2016 6th international conference - cloud system and big data engineering (confluence)* (pp. 445–450). http://dx.doi.org/10.1109/confluence.2016.7508161.

Mandler, G. (2002). Origins of the cognitive (r)evolution. *Journal of the History of the Behavioral Sciences*, *38*(4), 339–353. http://dx.doi.org/10.1002/jhbs.10066.

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, *503*(7474), 78–84. http://dx.doi.org/10.1038/nature12742.

Marder, E., Abbott, L. F., Turrigiano, G. G., Liu, Z., & Golowasch, J. (1996). Memory from the dynamics of intrinsic membrane currents. *Proceedings of the National Academy of Sciences*, *93*(24), 13481–13486. http://dx.doi.org/10.1073/pnas.93.24.13481.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., et al. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289. http://dx.doi.org/10.1038/s41593-018-0209-y.

Mayor, J., Gomez, P., Chang, F., & Lupyan, G. (2014). Connectionism coming of age: Legacy and future challenges. *Frontiers in Psychology*, *5*, 187. http://dx.doi.org/10.3389/fpsyg.2014.00187.

McFarland, J. L., & Fuchs, A. F. (1992). Discharge patterns in nucleus prepositus hypoglossi and adjacent medial vestibular nucleus during horizontal eye movement in behaving macaques. *Journal of Neurophysiology*, *68*(1), 319–332. http://dx.doi.org/10.1152/jn.1992.68.1.319.

Merel, J., Botvinick, M., & Wayne, G. (2019). Hierarchical motor control in mammals and machines. *Nature Communications*, *10*(1), 5489. http://dx.doi.org/10.1038/s41467-019-13239-6.

Miconi, T. (2017). Biologically plausible learning in recurrent neural networks reproduces neural dynamics observed during cognitive tasks. *ELife*, *6*, Article e20899. http://dx.doi.org/10.7554/elife.20899.

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, *16*(16), 5154–5167. http://dx.doi.org/10.1523/jneurosci.16-16-05154.1996.

Millidge, B., Tschantz, A., & Buckley, C. L. (2020). Predictive coding approximates backprop along arbitrary computation graphs. ArXiv.

Minsky, M., & Papert, S. A. (1969). *Perceptrons: An introduction to computational geometry*. MIT Press.

Misman, M. F., Samah, A. A., Ezudin, F. A., Majid, H. A., Shah, Z. A., Hashim, H., et al. (2019). Classification of adults with autism spectrum disorder using deep neural network. In *2019 1st international conference on artificial intelligence and data sciences (vol. 00)* (pp. 29–34). http://dx.doi.org/10.1109/aidas47888.2019.8970823.

Miyawaki, Y., Uchida, H., Yamashita, O., Sato, M., Morito, Y., Tanabe, H. C., et al. (2008). Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron*, *60*(5), 915–929. http://dx.doi.org/10.1016/j.neuron.2008.11.004.

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, *319*(5869), 1543–1546. http://dx.doi.org/10.1126/science.1150769.

Moravčík, M., Schmid, M., Burch, N., Lisý, V., Morrill, D., Bard, N., et al. (2017). DeepStack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, *356*(6337), Article eaam6960. http://dx.doi.org/10.1126/science.aam6960.

Mozer, M. C. (1987). Early parallel processing in reading: A connectionist approach. In M. Coltheart (Ed.), *Attention and performance 12: The psychology of reading*. Lawrence Erlbaum Associates, Inc.

Mukamel, E. A., Nimmerjahn, A., & Schnitzer, M. J. (2009). Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron*, *63*(6), 747–760. http://dx.doi.org/10.1016/j.neuron.2009.08.009.

Musall, S., Urai, A. E., Sussillo, D., & Churchland, A. K. (2019). Harnessing behavioral diversity to understand neural computations for cognition. *Current Opinion in Neurobiology*, *58*, 229–238. http://dx.doi.org/10.1016/j.conb.2019.09.011.

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature protocols*, *14*(7), 2152–2176. http://dx.doi.org/10.1038/s41596-019-0176-0.

Nilsson, S. R., Goodwin, N. L., Choong, J. J., Hwang, S., Wright, H. R., Norville, Z. C., et al. (2020). Simple Behavioral Analysis (SimBA) – an open source toolkit for computer classification of complex social behaviors in experimental animals. http://dx.doi.org/10.1101/2020.04.19.049452, BioRxiv, 2020.04.19.049452.

Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. L. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology*, *21*(19), 1641–1646. http://dx.doi.org/10.1016/j.cub.2011.08.031.

Nonaka, S., Majima, K., Aoki, S. C., & Kamitani, Y. (2020). Brain hierarchy score: Which deep neural networks are hierarchically brain-like? http://dx.doi.org/10.1101/2020.07.22.216713, BioRxiv, 2020.07.22.216713.

Noudoost, B., & Moore, T. (2011). Control of visual cortical signals by prefrontal dopamine. *Nature*, *474*(7351), 372–375. http://dx.doi.org/10.1038/nature09995.

Oh, J., Oh, B.-L., Lee, K.-U., Chae, J.-H., & Yun, K. (2020). Identifying schizophrenia using structural MRI with a deep learning algorithm. *Frontiers in Psychiatry*, *11*, 16. http://dx.doi.org/10.3389/fpsyt.2020.00016.

O'Reilly, R., Braver, T., & Cohen, J. D. (1999). A biologically based computational model of working memory. In A. Miyake, & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 375–411). Cambridge University Press, http://dx.doi.org/10.1017/cbo9781139174909.014.

Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neuroscience & Biobehavioral Reviews*, 36(4), 1140–1152. http://dx.doi.org/10.1016/j.neubiorev.2012.01.004.

Parsapoor, M. (2016). Brain emotional learning-based prediction model (for long-term chaotic prediction applications). ArXiv.

Pereira, T. D., Aldarondo, D. E., Willmore, L., Kislin, M., Wang, S. S.-H., Murthy, M., et al. (2019). Fast animal pose estimation using deep neural networks. *Nature Methods*, 16(1), 117–125. http://dx.doi.org/10.1038/s41592-018-0234-5.

Petersen, A., Simon, N., & Witten, D. (2017). SCALPEL: Extracting neurons from calcium imaging data. ArXiv.

Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., & Scanziani, M. (2013). Inhibition of inhibition in visual cortex: The logic of connections between molecularly distinct interneurons. *Nature Neuroscience*, 16(8), 1068–1076. http://dx.doi.org/10.1038/nn.3446.

Pfeiffer, M., & Pfeil, T. (2018). Deep learning with spiking neurons: Opportunities and challenges. *Frontiers in Neuroscience*, 12, 774. http://dx.doi.org/10.3389/fnins.2018.00774.

Plis, S. M., Hjelm, D. R., Salakhutdinov, R., Allen, E. A., Bockholt, H. J., Long, J. D., et al. (2014). Deep learning for neuroimaging: A validation study. *Frontiers in Neuroscience*, 8, 229. http://dx.doi.org/10.3389/fnins.2014.00229.

Pnevmatikakis, E. A. (2019). Analysis pipelines for calcium imaging data. *Current Opinion in Neurobiology*, 55, 15–21. http://dx.doi.org/10.1016/j.conb.2018.11.004.

Pnevmatikakis, E. A., Soudry, D., Gao, Y., Machado, T. A., Merel, J., Pfau, D., et al. (2016). Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron*, 89(2), 285–299. http://dx.doi.org/10.1016/j.neuron.2015.11.037.

Prut, Y., Vaadia, E., Bergman, H., Haalman, I., Slovin, H., & Abeles, M. (1998). Spatiotemporal structure of cortical activity: Properties and behavioral relevance. *Journal of Neurophysiology*, 79(6), 2857–2874. http://dx.doi.org/10.1152/jn.1998.79.6.2857.

Pulvermüller, F., Tomasello, R., Henningsen-Schomers, M. R., & Wennekers, T. (2021). Biological constraints on neural network models of cognitive function. *Nature Reviews Neuroscience*, 22(8), 488–502. http://dx.doi.org/10.1038/s41583-021-00473-5.

Raj, S., & Masood, S. (2020). Analysis and detection of autism spectrum disorder using machine learning techniques. *Procedia Computer Science*, 167, 994–1004. http://dx.doi.org/10.1016/j.procs.2020.03.399.

Rajan, K., Harvey, C. D., & Tank, D. W. (2016). Recurrent network models of sequence generation and memory. *Neuron*, 90(1), 128–142. http://dx.doi.org/10.1016/j.neuron.2016.02.009.

Richards, B. A., Lillicrap, T. P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., et al. (2019). A deep learning framework for neuroscience. *Nature Neuroscience*, 22(11), 1761–1770. http://dx.doi.org/10.1038/s41593-019-0520-2.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11), 1019–1025. http://dx.doi.org/10.1038/14819.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408. http://dx.doi.org/10.1037/h0042519.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. http://dx.doi.org/10.1038/323533a0.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing. Volume 2: Psychological and biological models*. MIT Press.

Ruppin, E., Reggia, J. A., & Horn, D. (1996). Pathogenesis of schizophrenic delusions and hallucinations: A neural model. *Schizophrenia Bulletin*, 22(1), 105–121. http://dx.doi.org/10.1093/schbul/22.1.105.

Salvatori, T., Song, Y., Lukasiewicz, T., Bogacz, R., & Xu, Z. (2021). Predictive coding can do exact backpropagation on any neural network. ArXiv.

Sandberg, A., Tegnér, J., & Lansner, A. (2003). A working memory model based on fast Hebbian learning. *Network (Bristol, England)*, 14(4), 789–802.

Scellier, B., & Bengio, Y. (2017). Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Frontiers in Computational Neuroscience*, 11, 24. http://dx.doi.org/10.3389/fncom.2017.00024.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117. http://dx.doi.org/10.1016/j.neunet.2014.09.003.

Schoenberg, P. L. A., & David, A. S. (2014). Biofeedback for psychiatric disorders: A systematic review. *Applied Psychophysiology and Biofeedback*, 39(2), 109–135. http://dx.doi.org/10.1007/s10484-014-9246-9.

Schrimpf, M., Kubilius, J., Hong, H., Majaj, N. J., Rajalingham, R., Issa, E. B., et al. (2020). Brain-score: Which artificial neural network for object recognition is most brain-like? http://dx.doi.org/10.1101/407007, BioRxiv, 407007.

Shen, G., Horikawa, T., Majima, K., & Kamitani, Y. (2019). Deep image reconstruction from human brain activity. *PLoS Computational Biology*, 15(1), Article e1006633. http://dx.doi.org/10.1371/journal.pcbi.1006633.

Shimamura, A. P. (1995). Memory and the prefrontal cortex. *Annals of the New York Academy of Sciences*, 769(1), 151–160. http://dx.doi.org/10.1111/j.1749-6632.1995.tb38136.x.

Shmueli, G. (2010). To explain or to predict? *Statistical Science, 3*(25), 289–310.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484–489. http://dx.doi.org/10.1038/nature16961.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., et al. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, 362(6419), 1140–1144. http://dx.doi.org/10.1126/science.aar6404.

Soleimanitaleb, Z., Keyvanrad, M. A., & Jafari, A. (2019). Object tracking methods: A review. In *2019 9th international conference on computer and knowledge engineering (vol. 00)* (pp. 282–288). http://dx.doi.org/10.1109/iccke48569.2019.8964761.

Soltanian-Zadeh, S., Sahingur, K., Blau, S., Gong, Y., & Farsiu, S. (2019). Fast and robust active neuron segmentation in two-photon calcium imaging using spatiotemporal deep learning. *Proceedings of the National Academy of Sciences*, 116(17), Article 201812995. http://dx.doi.org/10.1073/pnas.1812995116.

Song, H. F., Yang, G. R., & Wang, X.-J. (2016). Training excitatory-inhibitory recurrent neural networks for cognitive tasks: A simple and flexible framework. *PLoS Computational Biology*, 12(2), Article e1004792. http://dx.doi.org/10.1371/journal.pcbi.1004792.

Stevens, J. R. (1992). Abnormal reinnervation as a basis for schizophrenia: A hypothesis. *Archives of General Psychiatry*, 49(3), 238–243. http://dx.doi.org/10.1001/archpsyc.1992.01820030070009.

Sturman, O., von Ziegler, L., Schläppi, C., Akyol, F., Privitera, M., Slominski, D., et al. (2020). Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology*, 45(11), 1942–1952. http://dx.doi.org/10.1038/s41386-020-0776-y.

Sun, J., Cao, R., Zhou, M., Hussain, W., Wang, B., Xue, J., et al. (2021). A hybrid deep neural network for classification of schizophrenia using EEG Data. *Scientific Reports*, 11(1), 4706. http://dx.doi.org/10.1038/s41598-021-83350-6.

Sussillo, D., & Abbott, L. F. (2009). Generating coherent patterns of activity from chaotic neural networks. *Neuron*, 63(4), 544–557. http://dx.doi.org/10.1016/j.neuron.2009.07.018.

Taherkhani, A., Belatreche, A., Li, Y., Cosma, G., Maguire, L. P., & McGinnity, T. M. (2020). A review of learning in biologically plausible spiking neural networks. *Neural Networks*, 122, 253–272. http://dx.doi.org/10.1016/j.neunet.2019.09.036.

Tang, J., Yuan, F., Shen, X., Wang, Z., Rao, M., He, Y., et al. (2019). Bridging biological and artificial neural networks with emerging neuromorphic devices: Fundamentals, progress, and challenges. *Advanced Materials*, 31(49), Article 1902761. http://dx.doi.org/10.1002/adma.201902761.

Thomas, A. W., Heekeren, H. R., Müller, K.-R., & Samek, W. (2019). Analyzing neuroimaging data through recurrent deep learning models. *Frontiers in Neuroscience*, 13, 1321. http://dx.doi.org/10.3389/fnins.2019.01321.

Thomas, M. S. C., Knowland, V. C., & Karmiloff-Smith, A. (2011). Mechanisms of developmental regression in autism and the broader phenotype: A neural network modeling approach. *Psychological Review*, 118(4), 637–654. http://dx.doi.org/10.1037/a0025234.

Todorov, M. I., Paetzold, J. C., Schoppe, O., Tetteh, G., Shit, S., Efremov, V., et al. (2020). Machine learning analysis of whole mouse brain vasculature. *Nature Methods*, 17(4), 442–449. http://dx.doi.org/10.1038/s41592-020-0792-1.

Ullman, S. (2019). Using neuroscience to develop artificial intelligence. *Science*, 363(6428), 692–693. http://dx.doi.org/10.1126/science.aau6595.

Uyulan, C., Ergüzel, T. T., Unubol, H., Cebi, M., Sayar, G. H., Asad, M. N., et al. (2020). Major depressive disorder classification based on different convolutional neural network models: Deep learning approach. *Clinical EEG and Neuroscience*, 52(1), 38–51. http://dx.doi.org/10.1177/1550059420916634.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. ArXiv.

Verschae, R., & Ruiz-del-Solar, J. (2015). Object detection: Current and future directions. *Frontiers in Robotics and AI*, 2, 29. http://dx.doi.org/10.3389/frobt.2015.00029.

Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dudzik, A., Chung, J., et al. (2019). Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, 575(7782), 350–354. http://dx.doi.org/10.1038/s41586-019-1724-z.

von Ziegler, L., Sturman, O., & Bohacek, J. (2021). Big behavior: Challenges and opportunities in a new era of deep behavior profiling. *Neuropsychopharmacology*, 46(1), 33–44. http://dx.doi.org/10.1038/s41386-020-0751-7.

Wakefield, J. C. (2016). Diagnostic issues and controversies in DSM-5: Return of the false positives problem. *Annual Review of Clinical Psychology*, 12(1), 1–28. http://dx.doi.org/10.1146/annurev-clinpsy-032814-112800.

Wasmuht, D. F., Spaak, E., Buschman, T. J., Miller, E. K., & Stokes, M. G. (2018). Intrinsic neuronal dynamics predict distinct functional roles during working memory. *Nature Communications*, 9(1), 3499. http://dx.doi.org/10.1038/s41467-018-05961-4.

Watanabe, T., & Niki, H. (1985). Hippocampal unit activity and delayed response in the monkey. *Brain Research*, 325(1–2), 241–254. http://dx.doi.org/10.1016/0006-8993(85)90320-8.

Weisenburger, S., & Vaziri, A. (2016). A guide to emerging technologies for large-scale and whole-brain optical imaging of neuronal activity. *Annual Review of Neuroscience*, 41(1), 1–22. http://dx.doi.org/10.1146/annurev-neuro-072116-031458.

Whittington, J. C. R., & Bogacz, R. (2019). Theories of error back-propagation in the brain. *Trends in Cognitive Sciences*, 23(3), 235–250. http://dx.doi.org/10.1016/j.tics.2018.12.005.

Wiltschko, A. B., Johnson, M. J., Iurilli, G., Peterson, R. E., Katon, J. M., Pashkovski, S. L., et al. (2015). Mapping sub-second structure in mouse behavior. *Neuron*, 88(6), 1121–1135. http://dx.doi.org/10.1016/j.neuron.2015.11.031.

Wu, Q., McGinnity, T. M., Maguire, L., Cai, J., & Valderrama-Gonzalez, G. D. (2008). Motion detection using spiking neural network model. (pp. 76–83). http://dx.doi.org/10.1007/978-3-540-85984-0_10.

Xu, K., Su, H., Zhut, J., Guan, J.-S., & Zhangt, B. (2016). Neuron segmentation based on CNN with semi-supervised regularization. In *2016 IEEE conference on computer vision and pattern recognition workshops* (pp. 1324–1332). http://dx.doi.org/10.1109/cvprw.2016.167.

Xue, X., Halassa, M. M., & Chen, Z. S. (2021). Spiking recurrent neural networks represent task-relevant neural sequences in rule-dependent computation. http://dx.doi.org/10.1101/2021.01.21.427464, BioRxiv, 2021.01.21.427464.

Yamada, T., Hashimoto, R., Yahata, N., Ichikawa, N., Yoshihara, Y., Okamoto, Y., et al. (2017). Resting-state functional connectivity-based biomarkers and functional MRI-based neurofeedback for psychiatric disorders: A challenge for developing theranostic biomarkers. *International Journal of Neuropsychopharmacology*, 20(10), http://dx.doi.org/10.1093/ijnp/pyx059, pyx059-.

Yamashita, Y., & Tani, J. (2012). Spontaneous prediction error generation in schizophrenia. *PLoS One*, 7(5), Article e37843. http://dx.doi.org/10.1371/journal.pone.0037843.

Yamins, D. L. K., & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*, 19(3), 356–365. http://dx.doi.org/10.1038/nn.4244.

Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences*, 111(23), 8619–8624. http://dx.doi.org/10.1073/pnas.1403112111.

Yan, W., Calhoun, V., Song, M., Cui, Y., Yan, H., Liu, S., et al. (2019). Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site FMRI data. *EBioMedicine*, 47, 543–552. http://dx.doi.org/10.1016/j.ebiom.2019.08.023.

Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T., & Wang, X.-J. (2019). Task representations in neural networks trained to perform many cognitive tasks. *Nature Neuroscience*, 22(2), 297–306. http://dx.doi.org/10.1038/s41593-018-0310-2.

Yang, W., & Yuste, R. (2017). In vivo imaging of neural activity. *Nature Methods*, 14(4), 349–359. http://dx.doi.org/10.1038/nmeth.4230.

Yassin, W., Nakatani, H., Zhu, Y., Kojima, M., Owada, K., Kuwabara, H., et al. (2020). Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis. *Translational Psychiatry*, 10(1), 278. http://dx.doi.org/10.1038/s41398-020-00965-5.

Yin, B., Corradi, F., & Bohté, . S. M. (2020). Effective and efficient computation with multiple-timescale spiking recurrent neural networks. ArXiv.

Yoshihara, Y., Lisi, G., Yahata, N., Fujino, J., Matsumoto, Y., Miyata, J., et al. (2020). Overlapping but asymmetrical relationships between schizophrenia and autism revealed by brain connectivity. *Schizophrenia Bulletin*, 46(5), 1210–1218. http://dx.doi.org/10.1093/schbul/sbaa021.

Zeng, L.-L., Wang, H., Hu, P., Yang, B., Pu, W., Shen, H., et al. (2018). Multi-site diagnostic classification of schizophrenia using discriminant deep learning with functional connectivity MRI. *EBioMedicine*, 30, 74–85. http://dx.doi.org/10.1016/j.ebiom.2018.03.017.

Zhang, X.-Q., Jiang, R.-H., Fan, C.-X., Tong, T.-Y., Wang, T., & Huang, P.-C. (2021). Advances in deep learning methods for visual tracking: Literature review and fundamentals. *International Journal of Automation and Computing*, 18(3), 311–333. http://dx.doi.org/10.1007/s11633-020-1274-8.

Zhou, P., Resendez, S. L., Rodriguez-Romaguera, J., Jimenez, J. C., Neufeld, S. Q., Giovannucci, A., et al. (2018). Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *ELife*, 7, 3270. http://dx.doi.org/10.7554/elife.28728.

Zhou, Z., Wu, T.-C., Wang, B., Wang, H., Tu, X. M., & Feng, C. (2020). Machine learning methods in psychiatry: A brief introduction. *General Psychiatry*, 33(1), Article e100171. http://dx.doi.org/10.1136/gpsych-2019-100171.

Zhu, G., Jiang, B., Tong, L., Xie, Y., Zaharchuk, G., & Wintermark, M. (2019). Applications of deep learning to neuro-imaging techniques. *Frontiers in Neurology*, 10(869), http://dx.doi.org/10.3389/fneur.2019.00869.

Zipser, D. (1991). Recurrent network model of the neural mechanism of short-term active memory. *Neural Computation*, 3(2), 179–193. http://dx.doi.org/10.1162/neco.1991.3.2.179.