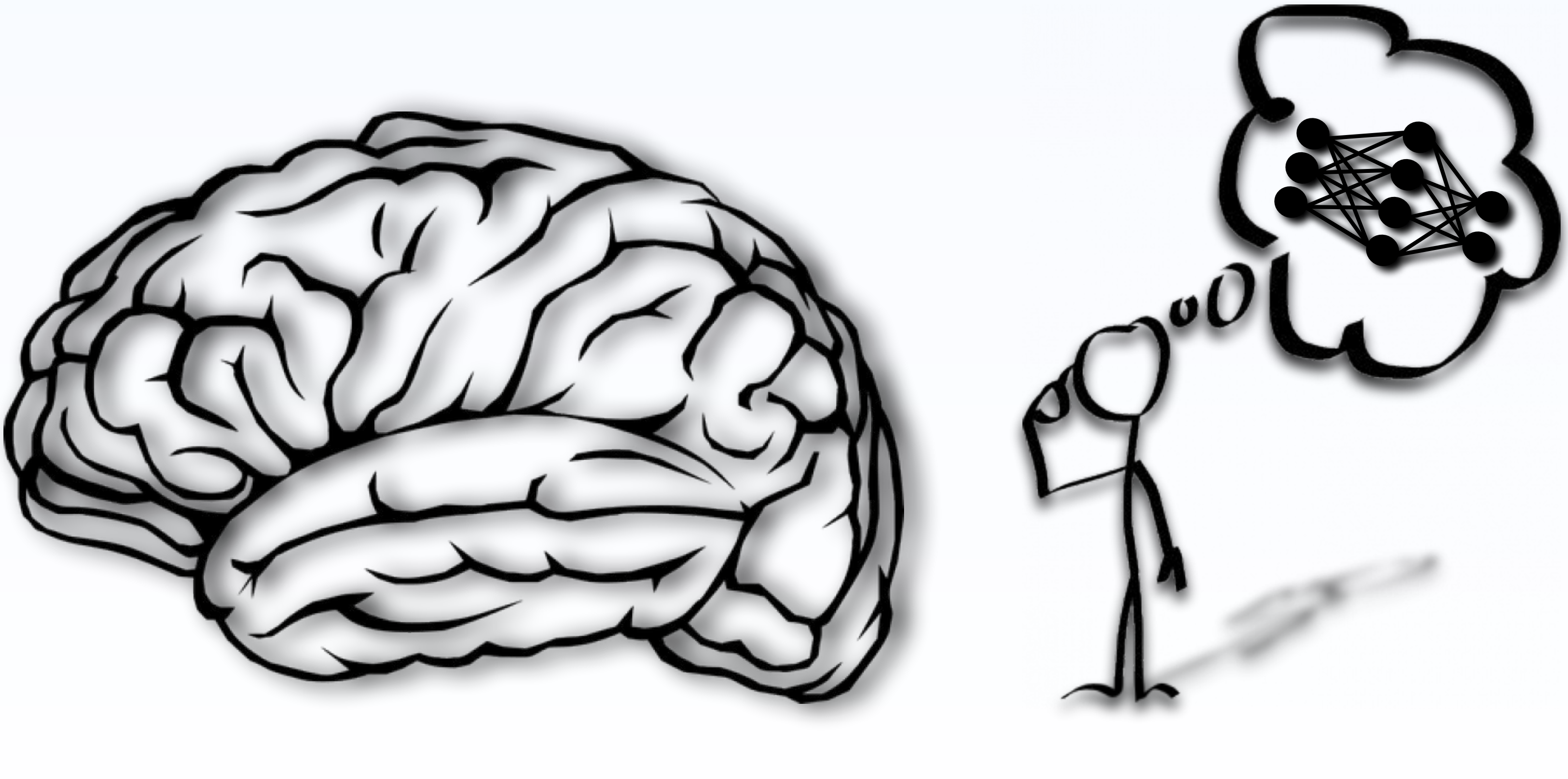
Biologically Constrained Spiking Neural Network for Image Classification



Nikos Malakasis

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
UNIVERSITY OF CRETE

Poirazi Lab
IMBB-FORTH

School of Medicine, University of Crete
Foundation for Research and Technology-Hellas
Institute of Molecular Biology and Biotechnology
Poirazi Lab

Biologically Constrained Spiking Neural Network for Image Classification

Nikos Malakasis

Thesis submitted
within the Master's Program in
Bioinformatics

Supervisor: Dr. Panayiota Poirazi

Three-Member Comittee:
Dr. Panayiota Poirazi
Dr. Pantelis Topalis
Dr. Pavlos Pavlidis

October 22, 2021

Thesis advisor: Dr. Panayiota Poirazi *Nikos Malakasis*

Biologically Constrained Spiking Neural Network for Image Classification

ABSTRACT

In the modern era of machine learning and artificial intelligence, artificial neural networks are widely used as the “brains” for smart software applications. Even though these algorithms are inspired by neurobiology, they lack features that neuronal circuits utilize to facilitate learning while employing learning rules that are not yet proven to be biologically plausible. In the presented study, we hypothesized that a biologically constrained neural network could perform image classification by utilizing network mechanisms and plasticity rules employed in the brain during learning. The network model includes an input layer and a single hidden layer that consists of multicompartmental spiking neurons with active dendrites, both excitatory pyramidals and inhibitory interneurons. Input neurons are sparsely connected to the hidden layer, and learning occurs by utilizing a class-specific learning signal and a synaptic tag-and-capture plasticity rule. We show that this biological network can perform binary image classification while maintaining the efficiency benefits of biological circuits. Moreover, this work highlights how structural plasticity improves performance by allowing the network to maximize its capacity. This model is arguably restricted due to computational inefficiency on regular hardware, which could be solved by implementing it on a neuromorphic platform. However, it can still serve as a template to test the advantages of biologically-inspired architectures when compared to deep learning algorithms, like its ability to learn continually.

Contents

Acknowledgements

I am kicking off this appreciation section by expressing my gratitude to Dr. Panayiota Poirazi for a handful of things, starting with the opportunity that she provided me with to work in her lab. It was an exciting educational year that I will hold on to for the rest of my life. Besides that, she is an excellent supervisor, definitely better than I could ever ask for. She is hard working, insightful, passionate about science and humble, constituting an ideal example of a scientist to look up to. Additionally, I am thankful for all of the advice that she gave me regarding my project. Without her counseling, this work would not have been realizable.

The second person that I am extremely thankful towards is Dr. Spyridon Chavlis, my co-supervisor and teacher. His teachings provided me with a strong educational foundation and his guidance in the lab with the appropriate blueprints to build this work from the ground up.

I would also like to thank all of the Poirazi lab members for the interesting, educational, sometimes meaningless, but always enjoyable conversations. Last but not least, I would like to thank my family, my girlfriend and all of the friends that supported me throughout this past year, either directly or indirectly.

1 Introduction

1.1 Motivation

Deep learning is a relatively new research field based on artificial neural networks, a family of machine learning algorithms inspired by the brain[1]. Despite being inspired by the brain, though, these algorithms usually utilize multiple biologically implausible mechanisms to improve their performance, which leads them further and further away from serving as a computational learning model for the brain, which is what we were interested in for this project. Thus, we decided to build a neural network model that is closer to biology than the ones mentioned above[1], as it could serve as a template on which brain functions related to learning can be evaluated and potentially unravel the advantages of the brain when it comes to learning and memory processes.

1.2 Background

In this section, background information will be provided regarding the required knowledge from the literature. Firstly, neuroscience-related information is presented, starting with details concerning the circuitry of the brain and its main units of computation and communication, moving on to knowledge regarding how learning occurs in the brain at the network and the cellular level and closing the neuroscience introductory part with computational neuroscience modeling. Then follows an introduction to machine learning, where different types of learning are discussed. Finally, the introductory chapter concludes with an overview of artificial neural network algorithms, inspirations, and constraints drawn from biology, and details on spiking neural networks, the subcategory of algorithms to which this work's network belongs.

1.2.1 Canonical circuitry of the brain

1.2.1.1 Neurons

Neurons are the cells that represent the fundamental computational units of the nervous system. A neuron can be divided into three major functional compartments: the dendrites that receive inputs and transmit them to the soma, the soma which is the body of the cell and its main computational compartment and the axon that transmits the output of the soma to postsynaptically connected neurons[2]. They can be classified into many different categories based on morphology, topology, or functionality. The scope of this subsection is on the functional division regarding excitation and inhibition in the brain. Excitatory neurons release the neurotransmitter glutamate, and inhibitory neurons release gamma-amino-butyric acid (GABA)[3]. Pyramidal neurons are the most dominant family in the brain among the neurons that have the former role. At the same time, interneurons are the most dominant family in the brain among the neurons that have the latter one.

1.2.1.1.1 Excitatory Pyramidal Neurons

Pyramidal neurons are a typical neuron class found in the cerebral cortex of every mammal, as well as in birds, fish and reptiles. They are also common in subcortical areas such as the hippocampus and the amygdala. They are named for their shape: typically, they have a soma that is shaped like a teardrop or rounded pyramid. They also tend to have a conical spray of longer dendrites that emerge from the pointy end of the soma (apical dendrites) and a cluster of shorter dendrites that emerge from the rounded end (basal dendrites)[3].

Pyramidal neurons comprise about two-thirds of all neurons in the mammalian cerebral cortex, which places them center-stage for many important cognitive processes[4]. What makes them special is their numerical dominance and the fact that they are often projecting their axons for long distances, sometimes out of the brain altogether. For example, pyramidal neurons in layer 5 of the motor cortex send their axons down the spinal cord to drive muscles. Thus, pyramidal neurons might be thought of as the “movers and shakers” of the brain[3].

They have a strong family resemblance, with their upright posture of apical and basal dendrites, but they vary in their appearance across different species and cortical regions. Molecular approaches are revealing genetic diversity amongst them, confirming the existence of distinct subtypes[3,5].

1.2.1.1.2 Inhibitory Interneurons

Although they constitute only 20-25% of all neurons in the cortex, inhibitory interneurons are very diverse, with different morphologies, sizes, intrinsic properties, connectivity patterns, and protein expression. A considerable effort has been made in recent years to classify them into subgroups based on their molecular properties. Two of the major subgroups into which they are classified are the parvalbumin expressing interneurons (PV) and the somatostatin expressing interneurons (SST)[6]. PV interneurons are the most represented cell type in the cortex. They have fast channels that grant them short recovery periods. They receive inputs from both excitatory neurons in the cortex as well as the thalamus. At least in layer 4, the major recipient layer of the cortex, they are the main source of fast feed-forward inhibition onto neighboring excitatory neurons, where, upon receiving inputs from the thalamus, they get activated and, in turn, inhibit their neighboring excitatory neurons in the forward direction[7,8]. Furthermore, they have been predicted to possess active dendrites, both supra and sublinear, depending on the dendrite’s volume[9]. SST interneurons are the second most abundant class of interneurons in the cortex, constituted mainly by Martinotti cells, located in layers 2-6, with density increasing with depth. They possess oval cell bodies and axons decorated with spiny boutons and beaded dendrites. A unique feature is that their axon ascends all the way to layer 1, spreading horizontally to neighboring columns and making synapses on the dendritic tufts of pyramidal neurons. These interneurons might be activated preferentially during periods of increased network activity in contrast to FS cells, which might be transiently activated preferentially when network activity is low[6,10,11].

1.2.1.2 Synapses

Neurons communicate through synapses, a phenomenal dynamic structure that permits the transmission of signals from the terminal of a presynaptic neuron's axon to the dendrites, soma, or axon of a postsynaptic one, or even to a different target effector cell[12]. They can be either electrical, where the signal is transmitted by special channels called gap junctions that are capable of passing an electric current, causing voltage changes in the presynaptic cell to induce voltage changes in the postsynaptic cell[13], or chemical, where the electrical presynaptic activity is converted to secretion of neurotransmitters via a canonical release machinery, while postsynaptic specialization senses those neurotransmitters via diverse receptors[14].

Synapses play a protagonistic role in learning and memory. Due to their dynamic nature, depending on specific rules, they are either strengthened or weakened, leading to memory formation. The release of neurotransmitters can be controlled on the presynaptic end, while the postsynaptic cell can alter the function and/or the number of receptors on its surface. In brief, this interplay between pre and postsynaptic cells in the synaptic cleft is known as synaptic plasticity. The strengthening of a synapse is known as long-term potentiation (LTP), whereas the weakening is known as long-term depression (LTD)[15]. Synaptic plasticity will be discussed further below, in **Section 1.2.2.**.

1.2.1.3 Dendrites

Dendrites, as stated above, are the compartment of the neuron that usually receives presynaptic input[2]. Their anatomy is quite variable between neuronal families and sometimes different even within the same family. Pyramidal neurons, for example, possess two distinct groups of dendrites, the basal and the apical ones. As implied by their name, the former emerge from the rounded basis of the soma and are short, while the latter emerges from the apex of the pyramidal cell and are way longer than the basal ones, forming a structure known as the dendritic tuft. Apical dendrites closer to the soma are labeled as the proximal ones, while those positioned closer to the end of the tuft are labeled as distal.

For years, dendrites have been considered as passive cellular antennas that practically receive inputs and transfer them to the single neuronal processing unit, the soma. However, multiple relatively recent modeling and experimental studies have proven that dendrites themselves are active neuronal subunits that process inputs before transmitting them to the soma[9,17–21].

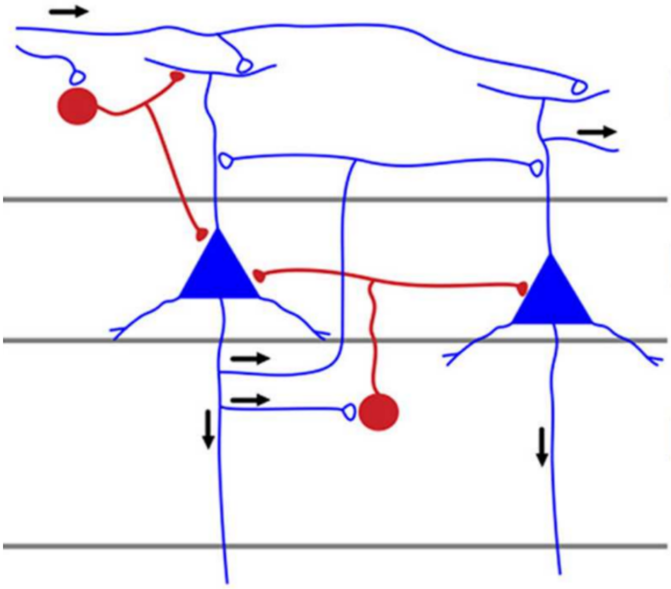


Figure 1.1. Simplified brain circuit. A brain circuit contains both excitatory (blue) and inhibitory (red) neurons. The neurons possess dendrites that receive inputs and communicate with other neurons via synapses. Adapted from Bosman and Aboitiz (2015)[16].

Dendritic nonlinearities can either be sublinear or supralinear. Sublinear integration of input is either mediated by passive [22,23], or by active dendritic properties[24,25], while supralinear integration is mediated only by active dendritic properties[20,21,26]. The importance of those nonlinearities has been shown in both computational modeling studies[9,17,18,27–29] and experimental ones[30–34].

Given that dendrites are the primary recipients of inputs, they have been strongly associated with learning and memory. Especially apical dendrites, since the role of their active properties in signal transduction to the soma adds an extra layer of complexity to the neural code that has even been used in various ways in brain-inspired deep learning algorithms[35–37].

1.2.2 Learning in the Brain

1.2.2.1 Memory Engrams

As mentioned above, learning in a neural circuit is facilitated by the dynamic interactions between neurons as they respond to stimuli. Eventually, those interactions lead to stable connections between a subset of neurons. Those neurons will respond to particular stimuli together and in a tuned manner, resulting in what is known as the memory of the stimuli. This subset of neurons is what is known as a memory engram[38,39].

An engram is a unit of cognitive information imprinted in a physical substance, theorized to be how memories are stored as biophysical or biochemical changes in the brain or other biological tissue response to external stimuli[40–42]. This hypothetical material basis of learned information was first conceived by Richard Semon, who theorized that learning induces persistent changes in specific brain cells that retain information and are subsequently reactivated upon appropriate retrieval conditions[38,43]. However, early experimental attempts to pinpoint localized engrams in rodent brains were inconclusive[44]. Experimental evidence for memory localization in the brain emerged from two separate observations by neurosurgeons years later. The first one was the induced recall of episodic memory by partial stimulation of the lateral temporal cortex[45]. In contrast, the second one was the observation that humans lacking a large part of the medial temporal lobe, including the hippocampus, showed amnesia for episodic memories[46]. Ever since the hippocampus has been identified as the brain region for contextual memory storage and retrieval by multiple studies in rodents[47,48], and besides the hippocampus, numerous brain regions have been reported as storage locations for different kinds of memories. As for identifying memory engrams *per se*, even though they remained elusive until recently, they have also been pinpointed in multiple studies[40−42].

1.2.2.2 Synaptic Plasticity

For memory engrams to be formed, changes in the synapses between neurons are required. Ramón y Cajal addressed this, proposing that synaptic strengthening should be a mechanism of memory storage[49]. However, the founder of the primary synaptic plasticity rules was none other than Donald Hebb, pointing out that neurons that fire together wire together[50]. His plasticity rules are still used in many different variations, forming a family of plasticity rules known as Hebbian rules.

These rules, however, refer to functional plasticity. Besides that kind of plasticity, other kinds of plasticity have also been characterized, like the structural and the homeostatic one, as discussed in the following subsections.

1.2.2.2.1 Functional Plasticity

In this subsection, the focus is on the plasticity that occurs in a defined synapse between two neurons. Depending on their correlated activity, this synapse is either strengthened or weakened for longer or shorter periods of time, depending on the type. This kind of plasticity has been modeled in many different ways. Two of the most known models are the standard Hebbian model and the spike timing-dependent plasticity (STDP) model. However, there are more functional synaptic plasticity models, like the one used in this work, the synaptic tag and capture model.

As already mentioned, the Hebbian model of plasticity is a pretty simple one, where neurons that fire together wire together. In practice, this means that if the presynaptic neuron fires in a short time window before the postsynaptic one, the synapse is strengthened via LTP[50]. This rule captures the essence of synaptic plasticity. Still, it leaves a lot of holes open, like the fact that it does not account for LTD, meaning that synapses are weakened only by homeostatic mechanisms. For that reason, many variations of it have emerged throughout the years that try to fill in those holes.

The most popular Hebbian rule variation is the STDP as mentioned earlier, model, which is widely used in computational neuroscience learning models as the synaptic plasticity mediator[51–56]. As a plasticity model, it is quite abstract, as the Hebbian one, but it fills in many holes that Hebb left open. First of all, besides LTP, which occurs when the presynaptic neuron fires in a time window before the postsynaptic one, this model accounts for LTD in an anti-Hebbian manner, meaning that if the presynaptic neuron fires in a time window after the postsynaptic one, the synaptic weight is weakened. One second important detail that it accounts for is that weight update is not all or none, but it scales in both directions, depending on the time interval between the pre and postsynaptic spikes[57].

While STDP covers how functional plasticity works abstractly, plasticity models have emerged that establish the causal relationship between synaptic plasticity and the biophysical, molecular machinery of a synapse. One of those is the synaptic tag and capture model[58,59]. In brief, LTP or LTD in this model requires both synaptic tags and plasticity-related proteins (PRPs) to occur. Synaptic tags are abstract entities, which signal that specific synapses should undergo synaptic plasticity and have not been explicitly identified at a molecular level. At the same time, PRPs are proteins that are involved in general with synaptic plasticity, like the AMPA receptor, CaMKII subunits, and cytoskeleton-related proteins like MAP2 and Arc60, which are either produced in the soma and then distributed to the synaptic sites, or they are even produced locally at dendrites[60,61]. This is also a model with many variations in terms of how tag and PRP production is induced. One of the most prevalent variations involves calcium ion influx as the trigger for tagging and protein synthesis[62], while calcium concentration determines whether the synapse is tagged for LTP or LTD[9,19].

1.2.2.2.2 Structural Plasticity

Structural plasticity is more straightforward than the functional one. It is defined in general as the ability of the brain to change its physical structure as a result of learning[65]. That refers to anything from neurogenesis to synaptic reallocation. This structural adaptability of the brain is a great asset. Besides coping with common events in the brain, it has been shown to be one of the mediators of functionality preservation after brain-damaging diseases or lesions[66–68]. Even though structural plasticity is easy to understand, there are many open questions regarding its occurrence. More specifically, during learning, synapses are pruned and reallocated consistently, while information is preserved. How synapses are reallocated with minimal or even no memory loss is still an open question.

1.2.2.2.3 Homeostatic Plasticity

Last but not least, homeostatic plasticity needs to be addressed. The aforementioned mechanisms, due to their dynamic nature, can drive a microcircuit to destabilization. For that particular reason, mechanisms that can modulate the stability of the

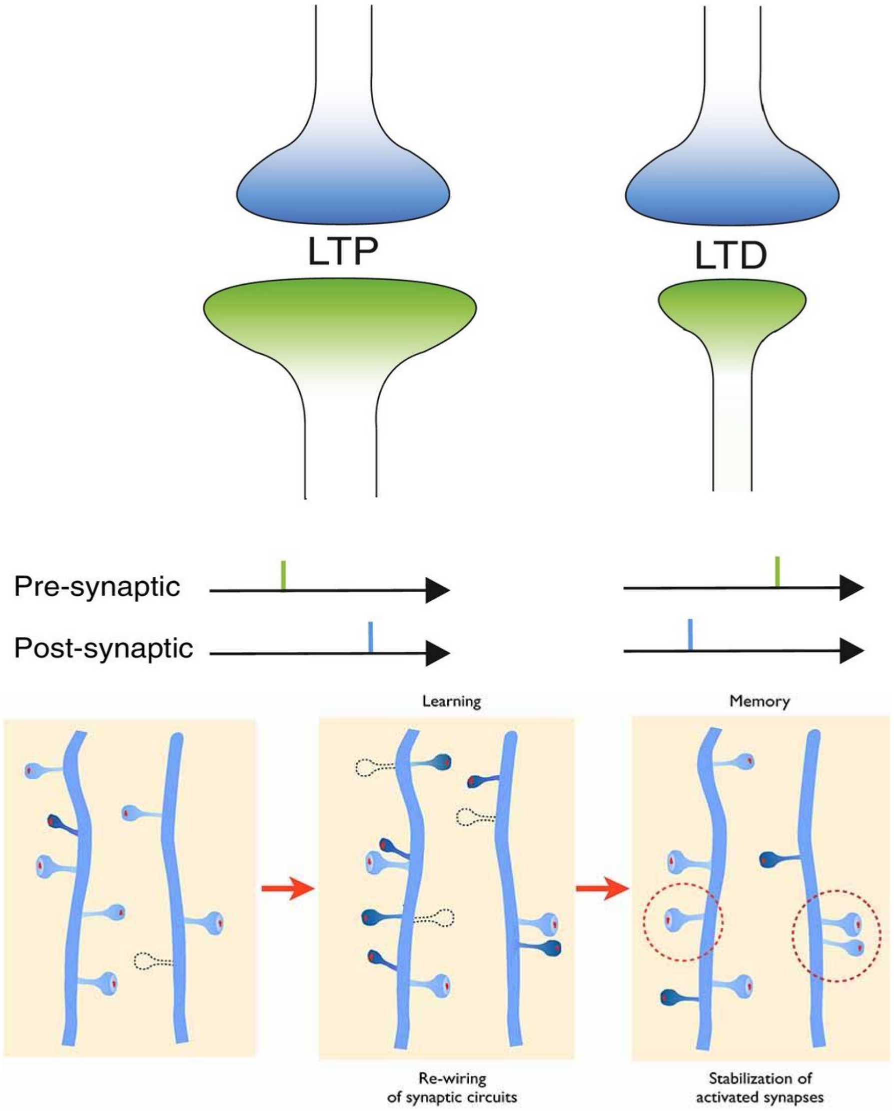


Figure 1.2. Synaptic plasticity. (top) An example of functional plasticity. Adapted from Andersen et al. (2017)[63]. **(bottom)** An example of structural plasticity. Unused small synapses are rewired, while active ones are stabilized. Adapted from Bernardinelli et al. (2014)[64].

brain network are needed.Quite a few mechanisms of homeostatic plasticity have been identified in brain networks, such as ion channel expression[69], co-regulated membrane conductances[70], neuromodulation[71], and extracellular protein regulation[72], that induce or reduce excitability or synaptic strength, depending on the microcircuit's state[73,74].

1.2.3 Machine Learning

Machine learning (ML) is the study of computer algorithms that can automatically improve through experience and data[75]. It is a rapidly evolving field, with algorithms that are commonly used in data science and artificial intelligence. With regards to how learning occurs in a machine learning algorithm, that algorithm can be classified as a supervised, unsupervised, or reinforcement learning algorithm.

1.2.3.1 Supervised Learning

Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs[76]. This mapping happens through a process where inputs are mapped to outputs via an initial, not ideal, function, generating an output that is compared to the optimal one, causing an error. The error is then used to optimize the initial function. This process is called training and the purpose of the function post-training is to map inputs to correct outputs that were not in the original training set. Examples of well-known supervised learning algorithms are support vector machines[77], k-nearest neighbor[78], and decision trees[79].

1.2.3.2 Unsupervised Learning

Unsupervised learning is a type of machine learning in which the algorithm is not provided with any pre-assigned labels or scores for the training data[80,81]. Thus, the algorithm has to learn how to solve the problem by observing the data and discovering potential patterns. Unsupervised algorithms' functionality is quite different from that of supervised ones, as supervised algorithms tend to perform better in classification tasks. However, they cannot operate that well without enough data, nor can they operate without labels, obviously. Unsupervised algorithms can detect patterns in data and group them based on those, known as clustering, or even determine the importance of certain features in a dataset and use that information to rank, combine or remove samples. Common examples of unsupervised learning include principal component analysis[82] and k-means clustering[83].

1.2.3.3 Neural Networks and Deep Learning

One of the primary reasons for machine learning's popularity during the last decade is the emergence of deep learning. Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised or unsupervised[1,84,85]. In this section, classic artificial neural networks (ANNs) and certain successful variations inspired by the brain will be discussed, as well as how biological constraints could and should, for certain purposes, pull them closer to biological plausibility. The section closes with the recent, most bioinspired variation of neural network algorithms, the spiking neural networks.

1.2.3.3.1 Artificial Neural Networks (ANNs)

ANNs are computational models inspired by how biological neural networks operate. There are many different variations of them, with different applications in classification, computer vision, natural language processing, artificial intelligence, and many more fields. They can be defined as directed graphs, where nodes are the artificial neurons, also known as perceptrons, and edges are the connections between them.

The neurons are organized into 3 different layers: the input layer that passes the information into the network, the hidden layers that process the information, and an output layer that generates predictions. In a typical ANN, all neurons in a layer are connected to every neuron in the next layer with a weight value that determines the sign and magnitude of the connection. The information is propagated forward in the network as a weighted sum of inputs plus a bias term, which is passed through an activation function in each neuron until it reaches the output layer, where a prediction is generated. Based on that predicted output and the desired output, during training, a loss value is generated. And thus, learning is an optimization problem in practice, where the goal is to minimize that loss. The optimization algorithm used to achieve that is known as (Stochastic) Gradient Descent. It is efficiently executed with the help of backpropagation-of-the-error (backprop), an automatic differentiation algorithm for calculating gradients with respect to parameters in a neural network. Based on those gradients, the network weights are updated, leading it to loss minimization and task learning[1,86].

From the time that they were first invented until today, Neural Network algorithms have been in a very close relationship with biology and specifically neuroscience[86–88]. The brain, being the most efficient and flexible learning model in existence, keeps on providing methods to improve ANNs in every possible aspect, leading them to great success over and over.

Probably the grandest successful example of neuro-inspiration in the modern neural network era are the convolutional neural networks (CNNs)[89]. Historically, the first convolutional neural network was published as the "neocognitron", a computational neuroscience model of the visual cortex[90], inspired by the work of Hubel and Wiesel on receptive fields[91]. Since then, CNNs evolved rapidly, resulting in a breakthrough performance in image recognition tasks when trained with gradient descent[89,92].

CNNs bear many advantages. Instead of fully connected layers, the layers perform convolutions. Typically, the convolution operation is the dot product of a sliding kernel with the layer's input matrix. This product is then passing through an activation function. As the convolution kernel slides along the input matrix, this operation generates a feature map, which, in turn, constitutes the input of the next layer. A noticeable advantage is the reduction of parameters needed for the network to operate, as convolutional layers are not dense, which reduces the amount of computation required per iteration by a lot. Since the amount of computations in terms of "width" is reduced, convolutional networks tend to take advantage of the extra computational power in "depth" meaning they tend to possess more layers than a standard fully connected ANN. Another advantage is interpretability since the generated feature maps per layer can be visualized, revealing the perceptive fields learned by the network.

There are more examples of how mechanisms found in the brain have been used to improve deep learning, like dropout, a mechanism inspired by synaptic pruning that prevents overfitting[93] or attention, a modern deep learning method utilized mainly by networks called transformers that is inspired by cognitive attention and yields state of the art performance in natural language processing[94].

1.2.3.3.2 Biological Constraints and Deep Learning Implausibilities

A common feature of the neuro-inspired examples in the previous subsection is that the biological mechanisms are embedded in a deep learning framework that is, however, biologically implausible for multiple reasons. Given that the brain is arguably the best learning model in existence, there is always merit in trying to imitate its functionality as much as possible. Even though neural networks draw inspiration from it, they are usually not constrained in many ways that the brain is. Those biological constraints could potentially benefit machine learning, as they may hide the key to the secret behind the brain's learning grandeur. Besides that, constrained models could also benefit neuroscience since predictions regarding neural mechanisms of cognition and learning stemming from these models are more credible than predictions coming from biologically implausible ones.

Standard deep learning algorithms violate multiple biological constraints. Firstly, neurons communicate with spikes, while ANNs propagate values. Those values could correspond to firing rates, but in that case, all the temporal information included in spike timings is lost. Moreover, biological neurons follow Dale's law[95], meaning that they can only release excitatory or inhibitory neurotransmitters, while ANN neurons have both excitatory and inhibitory postsynaptic connections. Additionally, ANNs are usually fully connected between layers, while biological networks are sparsely connected. The greatest inconsistency, though, between ANNs and the brain arises from training with backprop algorithm.

The aforementioned backprop algorithm via gradient descent is a mathematically efficient way of training a neural network while assigning credit where it's due. This credit assignment problem concerns determining how the success of a system's overall performance is due to the various contributions of the system's components. To solve it, this method assumes that error information is globally available throughout a network. A biological neuron, however, cannot possibly have global error information. Thus, the credit assignment problem should be solved in a local manner for a network to be biologically plausible. Local alternatives for the backprop in the brain have been proposed and

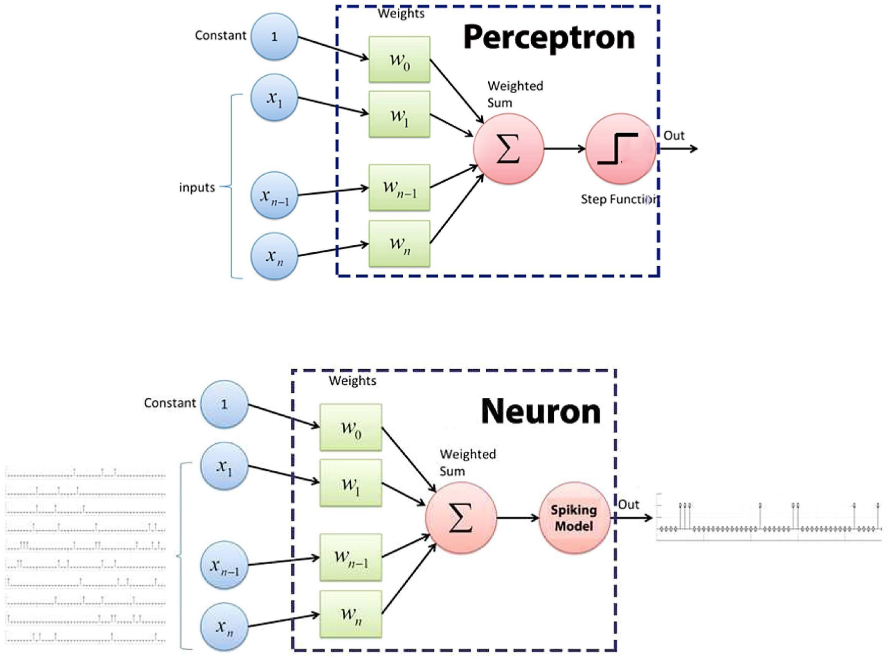


Figure 1.3. ANN Perceptron vs. SNN Neuron. A perceptron **(top)** receives and produces values as outputs, in an instantaneous manner, while a neuron **(bottom)** receives and produces spikes, in a temporal window that corresponds to a pattern presentation time. Adapted from Virgilio G et al. (2020)[99].

tested, most of which involve the electrophysiological segregation of input in the soma and the apical dendrites of a neuron, assuming that the apical dendrite receives feedback error signal. In contrast, the soma receives feed-forward information inputs[35,37].

Some additional inconsistencies regarding backprop and biology include that backprop uses the same weights for the forward pass and backward error propagation, while biological synapses are unidirectional, and also the fact that error propagation in backprop happens instantaneously, while in the brain, calculations and propagation require time[96]. Solutions to some of these problems have been suggested in a machine learning framework, like feedback alignment, where the weights in the backward pass are random and not the same as the ones in the forward one[97]. However, still, the framework itself is quite far away from the constrained biological reality. For that reason, a different but overlapping framework exists, where the networks consist of spiking neurons, and thus they are called spiking neural networks[98].

1.2.3.3.3 Spiking Neural Networks

Spiking neural networks (SNNs), also referred to as third-generation neural networks, are a variation of artificial neural networks. Instead of transmitting values instantaneously across space, imitate biological spiking neurons by transmitting spike trains for a certain amount of time per pattern presented[98]. These algorithms are closer to real neural circuits, which comes with both advantages, like efficiency in neuromorphic hardware, and disadvantages, primarily in performance.

SNN neurons can be modeled in various ways, as long as they communicate with spikes. Depending on the research goal, the neurons can be biophysically realistic, following the computationally extensive but accurate Hodgkin and Huxley neuron model[100], or they can even be as simple as the lightest in computations integrate and fire point neuron model[101]. In between those two extremes are the most commonly used alternatives like the Izhikevich neuronal model[102] or multiple variations of integrate-and-fire with adaptation and leakage mechanisms[103–107], since they permit larger circuit modeling while maintaining some biological validity.

Similarly, SNNs can be trained with multiple different learning rules. Biologically accurate models usually utilize synaptic plasticity mechanisms like the ones mentioned in **Section 1.2.2.2.**[108–110]. Conversely, gradient optimization that works well on ANN algorithms has been adapted to spiking networks by using surrogate gradients since spikes are not differentiable for efficiency purposes on neuromorphic hardware[111].

As previously stated, however, a great advantage of SNNs is that they can act as biological network models, on which theories regarding learning rules in the brain can and have been implemented. An example is the burst-dependent synaptic plasticity rule, which is based on the bursting behavior regulation of a neuron's soma by the NMDAR dependent activity of its apical dendrites. Specifically, assuming that feedback signals in a network target apical dendrites while feed-forward signals target basal ones, the coincidence between those two should lead to bursting and, according to the rule, LTP, while single spike activity leads to LTD[36].

To conclude, neural networks in general and spiking neural networks, in particular, are powerful tools in the hands of machine learning and neuroscience researchers. When implemented on the proper hardware, they allow for more efficient processes while also offering a more realistic modeling perspective regarding learning, memory, and cognition.

1.3 Research Purpose and Significance

This work lies in the intersection between machine learning and neuroscience. Its purpose is to use the machine learning framework for image classification to highlight the importance of certain brain network mechanisms in learning and memory through a biologically constrained spiking neural network model. Moreover, the network provided is quite unique. It incorporates multiple biophysical neuronal properties, abstractly modeled, constituting a prominent template for evaluating biologically inspired mechanisms *in silico* and their contribution to learning and memory. Finally, this is a model that could potentially unveil the basis behind the advantages of the human brain when compared to modern deep learning algorithms, like its ability to perform continual learning.

2 Methods

This network model was built on a previously published model for contextual memory formation and recall[9,19]. In this section, the properties of the model will be explained in thorough detail, starting from neuronal models and concluding with the details regarding the neuronal circuit.

2.1 Neuron Model

The network consists of both excitatory and inhibitory neuronal populations, both modeled as two-stage integrators[17], based on experimental data primarily from the hippocampal CA1 area. Neurons consist of a somatic spiking unit and independent dendritic subunits capable of nonlinear synaptic integration, dendritic spike initiation, and compartmentalized plasticity[18].

Synaptic input in each dendritic subunit is integrated according to **Equation 2.1**, where $t_{i,j}$ are the times of incoming spikes for synapse j, w_j is the weight of the synapse, τ_b is the voltage decay time constant, E_{syn} is the unitary EPSP, V_b is the dendritic branch voltage and $\delta(t)$ corresponds to Dirac delta.

$$\tau_b \frac{dV_b}{dt} = \sum_{i,j} w_j E_{syn} \delta(t - t_{i,j}) - V_b \tag{2.1}$$

Equation 2.1: Dendritic integration

The back-propagating action potential V_{bAP} (see **Equation 2.5**) is summed with V_b to determine the depolarization of the dendrite. When the sum of $V_b + V_{bAP}$, exceeds the dendritic spike generation threshold, θ_{dspike}, a dendritic spike is generated, which causes the voltage of the subunit, V_b, to rise instantaneously to V_{dspike}.

The synaptic input to the somatic neuronal compartment is calculated as shown in **Equation 2.2**, where g_{syn} corresponds to the dendritic coupling constant, $V_{b,n}$ to the voltage of the n_{th} branch, and $IPSC(t)$ to the inhibitory input that the neuron receives.

$$I_{syn}(t) = g_{syn} \sum_{n} (V_{b,n}(t)) - IPSC(t) \tag{2.2}$$

Equation 2.2: Total Synaptic Current reaching the Soma

The voltage response of the somatic subunit and its spiking output is modeled by an integrate-and-fire point unit with adaptation[112], as shown in **Equation 2.3** and

Equation 2.4. C is the somatic membrane capacitance, g_L is the leak conductance, E_L is the resting potential, g_{AHP} is the conductance of the afterhyperpolarization (AHP) current, and E_K is the AHP reversal potential. Somatic spiking and reset occur when the somatic voltage reaches a threshold V_T. **Equation 2.4** describes adaptive conductance g_{AHP} , where τ_{AHP} is the adaptation time constant, α_{AHP} is the quantal increase of g_{AHP} after a somatic spike which occurs at time t_{spike}, and $\delta(t)$ is the Dirac delta. The time constant τ_{AHP} can have two values that correspond to the neuron's high and low excitability levels.

$$C\frac{dV}{dt} = -g_L(V - E_L) - g_{AHP}(V - E_K) + I_{syn}(t) \tag{2.3}$$

Equation 2.3: Somatic Membrane Potential

$$\tau_{AHP}\frac{dg_{AHP}}{dt} = \alpha_{AHP}\delta(t - t_{spike}) - g_{AHP} \tag{2.4}$$

Equation 2.4: Adaptive Conductance

The backpropagating action potential is modeled by a depolarization component V_{bAP} which is added to all the dendritic subunits. $V_{bAP}(t)$ is modeled by an exponential (**Equation 2.5**) where E_{bAP} is the peak of the backpropagating depolarization and τ_{bAP} is the time constant of the backpropagating action potential.

$$V_{bAP}(t) = E_{bAP} * exp(-\frac{t}{\tau_{bAP}}) \tag{2.5}$$

Equation 2.5: Backpropagating Action Potential

2.2 Synaptic Plasticity Model

Learning in this model occurs according to the synaptic tag and capture synaptic plasticity model, which requires both synaptic tagging and the availability of PRPs for stable strengthening and weakening of synapses[59]. When a postsynaptic branch receives presynaptic inputs, it leads to calcium influx. At the end of a pattern presentation event, a certain amount of calcium ions have accumulated in the respective branches. Based on the individual calcium ion levels, a synaptic tag is generated that leads to long-term depression(LTD) of the corresponding synapse if the levels are low, or long term potentiation (LTP) if the levels are high. Once the synapses are tagged, plasticity-related proteins are produced that are consolidated together with the synaptic tags into the existing synaptic weights in the timespan of a couple of hours, leading to the synaptic weight update. This brief explanation of the synaptic plasticity mechanism will be dissected in the following subsections.

2.2.1 Calcium modeling

Calcium is the critical component of this plasticity mechanism, as it induces synaptic tags and PRP synthesis. The total calcium influx during the presentation of a pattern to a synapse determines the calcium C_{syn}, which models the calcium concentration near each synapse. Each incoming synaptic spike causes a step increase of calcium, which depends nonlinearly on the local depolarization of the dendritic branch where the synapse resides. We assume that calcium influx upon arrival of a presynaptic spike ΔC_{syn} is primarily through NMDA receptors[113], and is, thus, dependent sigmoidally on the dendritic membrane voltage, as described in **Equation 2.6**, where α_{Ca} is the maximum calcium ion influx and V corresponds to the sum of $V_b + V_{bAP}$.

$$\Delta C_{syn} = \alpha_{Ca} \frac{1}{1 + exp(-\frac{V - 30mV}{5mV})} \tag{2.6}$$

Equation 2.6: Calcium influx

2.2.2 Plasticity Related Proteins Production

It has been shown in multiple plasticity studies that both dendritic[114,115] and somatic[116,117] protein synthesis are essential for synaptic weight consolidation, separately or even both at the same time[118]. The previously published model[19] incorporated three different mechanisms for plasticity-related protein synthesis, where either somatic only, dendritic only, or both are happening. In the current study, only the somatic PRP synthesis model was used to generate the presented results. PRP synthesis initiation is modeled as an all-or-none phenomenon. When the calcium level, which is equal to the sum of all of its dendritic calcium levels, passes a threshold (P_{soma}), a PRP transient is generated. The PRP transient's time course is modeled after an alpha function, described by **Equation 2.7**. $H(t)$ corresponds to the Heaviside step function.

$$PRP_{soma}(t) = H(t - 20min)(\frac{t - 20min}{30min})exp(1 - \frac{t - 20min}{30min}) \tag{2.7}$$

Equation 2.7: PRP transient time course

2.2.3 Synaptic Tag Generation and Consolidation

As previously mentioned, the synaptic tag's strength and sign are determined according to the Calcium Control Model[119], thus lower levels of Ca^{2+} cause LTD, while higher levels cause LTP. This is described analytically in **Equation 2.8**. The synaptic tag does not alter the weight of the synapse immediately, but only after the capture of PRPs required for consolidation. Synaptic tags in this model decay exponentially with a time constant of 1 hour.

$$synTag(x) = \frac{1.3}{1 + exp(-10(10x - 3.5))} - \frac{0.3}{1 + exp(-19(10x - 2.0))} \tag{2.8}$$

Equation 2.8: Calcium dependent Synaptic Tag Generation

The synaptic tags are consolidated into synaptic weight over time, according to **Equation 2.9**. η is the network's learning rate, that can either be global or local per synapse, which will be discussed further in **Subsection 2.3.2.2**. *synaptictag* corresponds to the value of the synaptic tag that can either be positive or negative, α_s to the rate of synaptic tag consolidation and $PRP(t)$ to the PRP levels at time t. The weights are clipped in the range [0,1].

$$\Delta w = \eta * \alpha_s * synaptictag * PRP(t) \tag{2.9}$$

Equation 2.9: Weight update equation

2.2.4 Dynamic Excitability

Learning has been shown to increase the excitability of neurons participating in the formation of a given memory[120–125]. On the other hand, neurons with increased excitability are more likely to participate in the formation of a new memory engram[122,126,127]. The activation of transcription factor CREB (cAMP Response Element-Binding Protein) has also been found to modulate the excitability of neurons[128,129] through the reduction of the AHP current[122,130]. Therefore, it has been suggested that learning makes cells more amenable to be recruited in future learning events through the activation of CREB[121,131–134]. Finally, it has been proposed that CREB may also induce the downstream activation of its own repressors[121,122], which would reduce excitability after a certain period, thus creating a time window of increased neuronal activity excitability. Increased excitability is simulated through the transient reduction of the AHP current in the neurons in which PRP synthesis is triggered for approximately 12 hours after the learning event[19].

2.2.5 Homeostatic Plasticity

The effect of homeostasis on synaptic weights is modeled using a synaptic scaling rule[135]. According to this rule, the total synaptic weight of a model neuron remains constant. The synaptic weights w_j of each synapse are normalized according to **Equation 2.10**, where w_{init} is the initial synapse weight, N_{syn} the total number of synapses in the model neuron and τ_H the time constant of homeostatic synaptic scaling[19].

$$\tau_H \frac{dw_j}{dt} = 1 - \frac{\sum_j w_j}{w_{init} N_{syn}} \tag{2.10}$$

Equation 2.10: Homeostatic Synaptic Scaling

2.3 The Network

In the following section, the project's network model shall be dissected in thorough detail. Firstly, the network's structure and afterwards the bio-inspired rules that were implemented and tested to yield the results shown in the **Results** chapter will be explained.

2.3.1 Network Model and Structure

In brief, the network consists of one input layer and a hidden/output layer. The input layer is connected to the hidden one sparsely, and learning occurs with the help of a class-specific learning signal during training. In the testing phase, the learning signal is no more, and the output evaluation occurs based on a majority rule.

2.3.1.1 Input Translation and Transmission

First of all, the dataset that was used to conduct all of the presented experiments is the MNIST (Modified National Institute of Standards and Technology database)[136]. This well-known handwritten digit dataset is widely used for the benchmarking of machine learning models. It consists of 70,000, 28 × 28 pixel images of handwritten digits, from 0 to 9. The dataset is already split into a training set containing 60,000 images and a test set containing 10,000 images.

To use those images as an input for a spiking neural network, they need to be translated into spike trains. This was done by translating pixel intensity into firing frequency (**Equation 2.11**) and then generating a random spike train per pixel, with the corresponding firing frequency, for the pattern presentation time, resulting in $28 \cdot 28 = 784$ input neurons.

$$F_n = \frac{F_{max} * PI_n}{PI_{max}} \tag{2.11}$$

Equation 2.11: Pixel Intensity to Firing Frequency translation

The translated input is then transmitted to the hidden layer in spikes, through synaptic connections, for 4 seconds. After an input has been transmitted, a consolidation period of 2 hours follows. Synaptic tags and plasticity-related proteins are produced and consolidated into synaptic weights, leading to the weight update. After this period ends and the weights are updated, the next pattern is presented.

2.3.1.2 Network Connectivity and Computations

As previously mentioned, the network consists of an input and a hidden layer. The hidden layer possesses both excitatory pyramidal neurons and inhibitory interneurons. Input neurons are sparsely and randomly connected to the pyramidals of the hidden layer, with initial weights pulled from a uniform distribution in the range [0.1, 0.2]. Those are the only plastic weights of the network.

For learning to occur, the network is equipped with a supervised mechanism in the form of a learning signal neuron specific to each different class. Each class-specific neuron targets a distinct subpopulation of pyramidal neurons in the hidden layer and fires with a high firing rate F_{ls}. This ensures that those neurons that receive a strong input both from the input layer and from the class-specific neuron will participate in the learning procedure of the presented pattern, leading to the formation of a memory engram. In this project, classification was limited to two classes at a time (binary). Thus, the learning signal neurons were only two, and so were the corresponding pyramidal subpopulations.

Besides pyramidals, interneurons are also represented in the hidden layer. The network possesses both somatostatin expressing (SST) interneurons and parvalbumin expressing (PV) interneurons. In the current modeling scope, there are two differences between those two types of neurons, one being the fact that SST interneurons target the dendrites of the postsynaptic neuron, while PV target the soma, and the other being in their nonlinearities, as SST have supralinear dendrites, since increasing their excitability works better in the model, while PV have sublinear.

There are two groups in terms of functional role, one controlling the network's activity, since it is quite excitable, and another performing feedback inhibition. The former one consists of both SST and PV interneurons that are randomly connected pre and post synaptically with pyramidals in the hidden layer, while the latter consists of exclusively SST interneurons that are grouped in the same manner as pyramidal subpopulations are, receive input from one of those subpopulations each, and inhibit the rest of them.

2.3.1.3 Network Training and Testing

An iteration of the training loop has been fully described in the previous subsections. In order for the network to learn, multiple of those are needed, and training stops according to a stopping criterion that will be described in **Subsection 2.3.2.3**. Stopping according to that criterion occurs after approximately 100 to 200 iterations in the binary classification framework, given the hyperparameters used and depending on the digit pair, which corresponds to approximately 50 to 100 images per class.

After training is finished, the model needs to be evaluated in its ability to classify the patterns that it was presented with into their respective classes. To do that, the plastic weights are frozen, and the hidden layer receives input only from the input layer. After every testing pattern presented, the network's classification decision is determined by a majority rule. In order for the majority rule to be explained, the definition of active neurons is necessary. An active neuron in the hidden layer is a pyramidal neuron that fires more than 20 spikes during the pattern presentation time. Thus, the majority

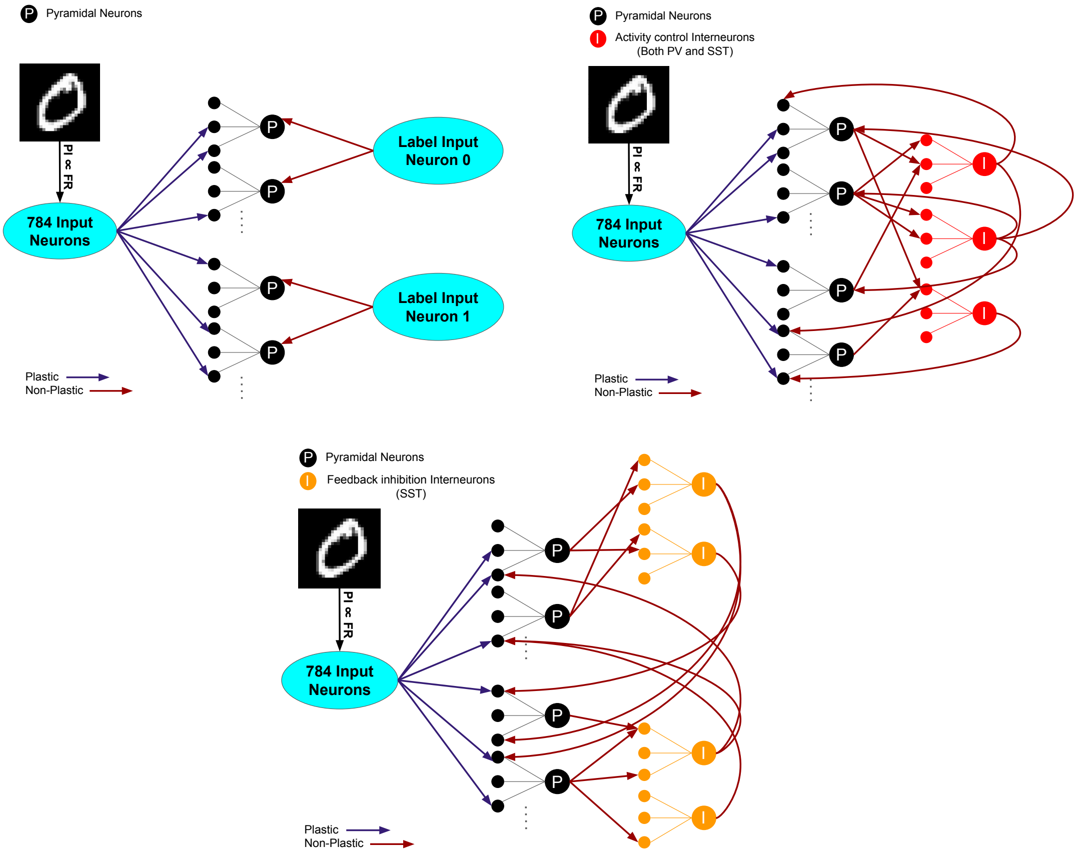


Figure 2.1. **Network structure and connectivity.**The network's input neurons are sparsely connected to the hidden layer and the hidden layer ones in turn to interneurons.**(top left)**Learning occurs with the supervision of a class specific neuron per subpopulation.**(top right)**The network's excitability is controlled by an activity control interneuron group, consisting of both PV and SST interneurons.**(bottom)**Feedback inhibition SST interneurons also support the learning procedure, by inhibiting all subpopulations, except the one that activates them.

rule compares the number of active neurons in each subpopulation. The class that corresponds to the subpopulation with the most active neurons is the network's decision.

2.3.2 Biologically Inspired Functions

In this subsection, as the subtitle suggests, biologically inspired functions that were added to the network will be discussed. The network has many biologically inspired elements, and quite a few of them were explained in previous subsections. The ones that will be explained below are functions that were added to improve the network's performance in classification, and experiments have been conducted to show their importance, as presented in the **Results** chapter.

2.3.2.1 Synaptic Turnover

Synaptic turnover is a structural plasticity mechanism that involves synaptic pruning and synaptogenesis. It has been shown that this mechanism is strongly associated with learning in multiple areas of the brain[137–144]. Evidence also suggests that turnover usually occurs at smaller and less stable synapses during learning events[137,138,145], thus, increasing the network's learning capacity[138,146]. Synaptic turnover was implemented in this work by randomly reinitializing any synapse with a weight lower than the maximum initial one after every 20 training iterations.

2.3.2.2 Local Adaptive Learning Rate

Another feature of biological synapses is the fact that as the synaptic spines grow and become more stable, their growth rate drops[138,147–150]. This is incorporated in the network in the form of an adaptive synapse-specific learning rate that is dependent on the synaptic weight value(**Equation 2.12**). $\eta_j(w)$ corresponds to the learning rate of synapse j given the weight value and follows a reverse sigmoidal, where η_{min} and η_{max} are the minimum and maximum learning rate value, respectively, *slope* and *mid* denote the sigmoidal's slope (i.e., steepness) and midpoint, respectively.

$$\eta_j(w) = -\frac{\eta_{max} - \eta_{min}}{(1 + exp(slope * (w - mid))} + \eta_{max} \tag{2.12}$$

Equation 2.12: Local Adaptive Learning Rate

2.3.2.3 Biologically Inspired Stopping Criterion

As stated above, larger spines have been observed to be more stable than smaller ones, leading to the hypothesis that the large ones should play a role in memory formation[138,147,149]. Given that, a stopping criterion was implemented that terminates training when a percentage of synaptic spines becomes larger than a threshold value, assuming that in that case, memory has already been formed. The spine size is characterized by its synapse's learning rate.

2.4 Constant and Parameter Table

Here, a table containing the values of the parameters and constants of the model is presented.

Parameter	Description	Value
τ_b	Passive dendritic integration time constant	20 msec
E_{syn}	Maximum unitary EPSP	4.0 mV for excitatory neurons 3.0 mV for inhibitory neurons

Parameter	**Description**	**Value**
θ_{dspike}	Depolarization threshold for dendritic spiking	25 mV
V_{dspike}	Dendritic spike max depolarization	50 mV
E_L	Somatic leakage reversal potential	0 mV
θ_{soma}	Voltage threshold for somatic spikes	20 mV
g_{syn}	Dendritic coupling constant	180 nS inh 120 nS exc
τ_{soma}	Somatic time constant	30 msec for excitatory neurons 10 msec for inhibitory neurons
τ_{AHP}	Adaptation time constant of excitatory neurons	180msec (slow adapting) 110msec (fast adapting)
α_{AHP}	Adaptation conductance increase after a spike	0.18 nS
E_K	Adaptation current reversal potential	-10 mV
τ_{bAP}	Back propagating action potential time constant	17 msec
E_{bAP}	Back propagating action potential max amplitude	30 mV
α_{Ca}	Calcium influx rate	1.1 msec^{-1}
P_{soma}	Calcium level threshold for somatic Plasticity-Related Protein synthesis	18.0 (a. u.)
τ_{PRP}	Time constant for PRP decay	60 mins
τ_H	Time constant of homeostatic synaptic scaling	7 days
α_s	Rate of synaptic tag consolidation	6.7 mins
T_{CREB}	Duration of increased excitability	12.5 hours
N_n	Number of Neurons in the hidden layer	120 neurons
N_b	Number of Branches	10 branches
N_{pyr}	Number of Pyramidals	80 neurons (40 per subpopulation)
$N_{ctrlinh}$	Number of control interneurons	10 SST interneurons 10 PV interneurons

Parameter	**Description**	**Value**
N_{latinh}	Number of feedback inhibition interneurons	20 SST interneurons
N_{input}	Number of Input Neurons	784 neurons
$N_{inp \rightarrow hid}$	Number of total synapses from the input layer to the hidden layer pyramidals	1750 synapses
N_{labinp}	Number of class specific input neurons	2 neurons (1 per subpopulation)
$N_{pyr \rightarrow inhctrl}$	Number of total synapses from the hidden layer pyramidals to the hidden layer control interneurons	100 synapses to PV 500 synapses to SST
$N_{inhctrl \rightarrow pyr}$	Number of total synapses from the hidden layer control interneurons to the hidden layer pyramidals	200 synapses from PV 5000 synapses from SST
$N_{pyr \rightarrow inhlat}$	Number of total synapses from the hidden layer pyramidals to the hidden layer feedback inhibition interneurons	320 synapses (160 per subpopulation)
$N_{inhlat \rightarrow pyr}$	Number of total synapses from the hidden layer feedback inhibition interneurons to the hidden layer pyramidals	320 synapses (160 per subpopulation)
$N_{labinp \rightarrow pyr}$	Number of total synapses from the class specific input neuron to the hidden layer pyramidals	160 synapses (80 per subpopulation)
W_{min}	Minimum initial synaptic weight for plastic synapses	0.1 (a.u.)
W_{max}	Maximum initial synaptic weight for plastic synapses	0.2 (a.u.)
η_{min}	Minimum Learning Rate	0.001 (a.u.)
η_{max}	Maximum Learning Rate	0.01 (a.u.)
$slope$	Slope of learning rate's reverse sigmoidal	10 (a.u.)
mid	Weight value of the midpoint of learning rate's reverse sigmoidal	0.5 (a.u.)
$N_{turniters}$	Number of iterations to synaptic turnover event	20 iterations
$W_{turnover}$	Weight threshold value for synaptic turnover	0.2 (a.u.)

Parameter	Description	Value
P_{stop}	Percentage of synapses with large spines required for stopping criterion	50%
η_{large}	Learning rate value above which a synapse's spine is considered large	0.0055 (a.u.)
F_{max}	Maximum firing rate for an input layer neuron	25 Hz
F_{labinp}	Firing rate of class specific input neuron	40 Hz
N_{runs}	Number of simulations per digit pair during classification experiments	18 simulations

Table 2.1: Model parameters with short descriptions and values

2.5 Computational Resources and Data Analysis

The simulations for this project were run on the high-performance computing (HPC) cluster (rocksclusters.org) of the Poirazi Lab, which consists of 624 cores and 3,328 GB shared RAM operating under CentOS Linux distribution. Data analysis was performed in python 3 using the NumPy[151] and SciPy[152] libraries. The plots were generated using Seaborn[153] and Matplotlib[154] libraries.

3 Results

3.1 Result Outline

In this chapter, as the title suggests, this work's results shall be presented. They can be categorized into two different parts. The first one is the model's performance on binary classification, and the second one is the comparison of the model's classification performance with and without one of the bio-inspired mechanisms at a time to show their importance in the model and also to infer their importance in biological neuronal systems. All of the parameters and constant values that were used for the following experiments can be found in **Section 2.4**.

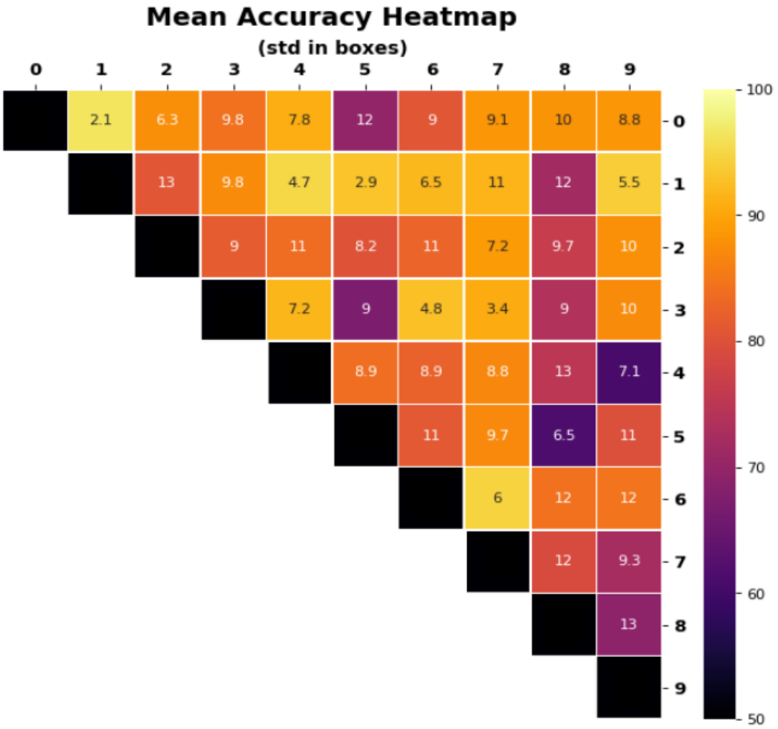


Figure 3.1. The model's baseline performance. Mean accuracy heatmap. Every square represents one digit pair. The color of each box corresponds to the mean classification accuracy for the particular pair across all of its simulations and the number inside to standard deviation.

3.2 Model Performance

The network is thoroughly explained in the **Methods** chapter. In brief, it includes an input layer and a single hidden layer. The hidden layer consists of multicompartmental spiking neurons with active dendrites, both excitatory pyramidal cells and inhibitory interneurons. Input neurons are sparsely connected to the hidden layer. Learning occurs by utilizing a class-specific learning signal and the synaptic tag-and-capture plasticity rule and structural plasticity. The synaptic learning rate is dependent on the weight value of each synapse, and learning stops once a certain percentage of synapses have a learning rate greater than a threshold value.

The network's ability to perform binary classification was tested for every pair of MNIST digits 18 times, and the results are shown in **Figure 3.1.** Even though it can classify some of the easier pairs consistently well, its accuracy drops for the harder ones and becomes inconsistent.

Regarding the drop in performance for the harder pairs, it can be assumed that it is a consequence of learning with only 100 to 200 images in total, meaning that it does not have enough experience to be able to tell apart small differences between very similar handwritten digits, like four and nine. Besides that, another potential reason is its size.

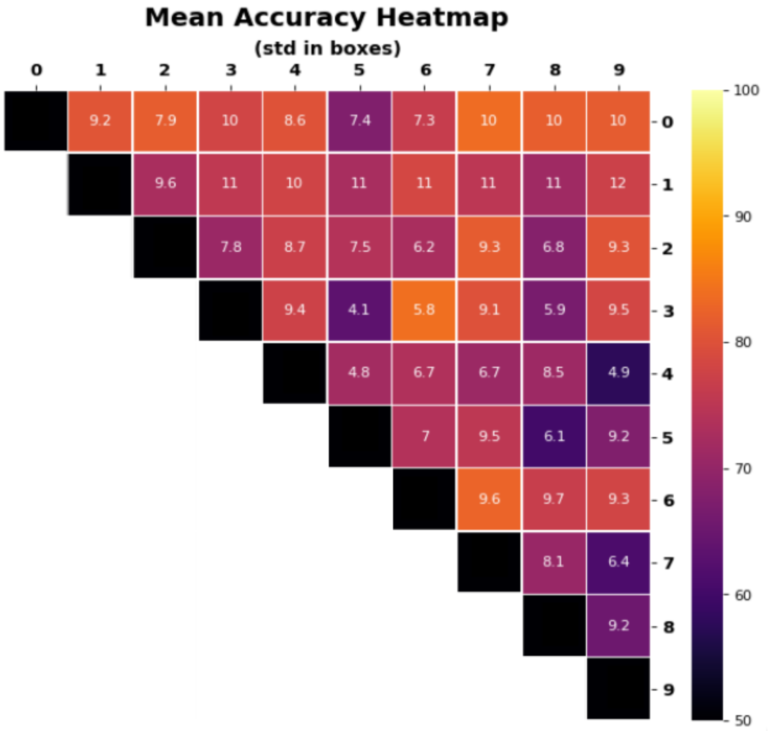
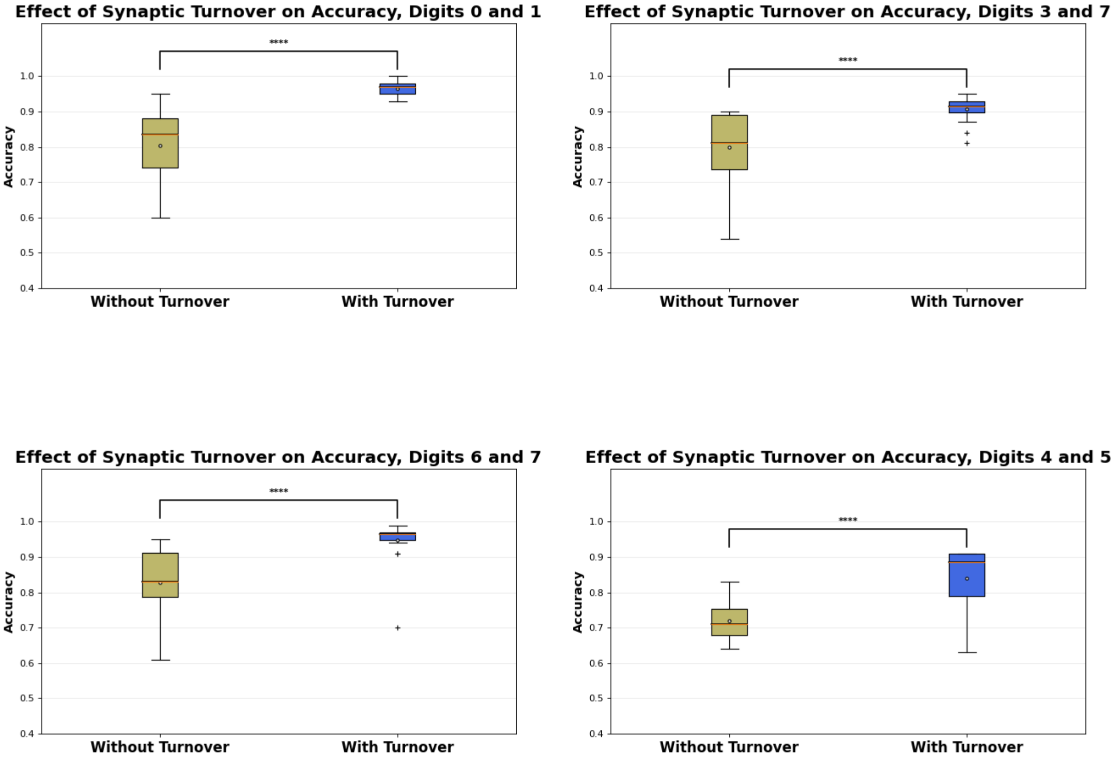


Figure 3.2. The model's performance without synaptic turnover. Mean accuracy heatmap. Every square represents one digit pair. The color of each box corresponds to the mean classification accuracy for the particular pair across all of its simulations and the number inside to standard deviation.

The same problems also apply for its inconsistency. For every different simulation of a particular pair of digits, the order of the training set changes, which means that it is trained with a different set of images each time. Another reason for the inconsistency lies in the sparseness of the network since for every different simulation of a particular pair of digits, the initial synapses are different, and even though synaptic turnover will eventually help with using every synapse of the network, the training process might be too short for it to do it.

Even so, this is not a network model built to outperform machine learning algorithms in classification. This network constitutes a spiking neural network template to test whether biological functions improve classification performance while staying as close to biology as possible with multiple constraints.


Effect of Synaptic Turnover on Accuracy, Digits 0 and 1

Effect of Synaptic Turnover on Accuracy, Digits 3 and 7

Effect of Synaptic Turnover on Accuracy, Digits 6 and 7

Effect of Synaptic Turnover on Accuracy, Digits 4 and 5

Figure 3.3. Comparison of no turnover network performance to baseline. Each boxplot represents the accuracy across 18 simulations for the particular digit pair in the given condition. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, Welch t-test.

3.3 Impact of Synaptic Turnover

Synaptic turnover is the bio-inspired mechanism with the most crucial role in performance out of the three shown in this chapter. Synapses that did not grow after twenty images were presented are pruned, and new ones take their place randomly in the network. This synaptic reinitialization allows the network to fully utilize its total capacity, resulting in greater efficiency and more consistent results.

That is observable in **Figure 3.2.**, where the network's performance without synaptic turnover is shown. Every pair is performing worse on average and less consistently, as indicated by the standard deviation. An explanation for that is that a sparse network without turnover is initialization dependent, and the probability for a small number of synapses to be allocated in a useful manner is low. The difference in mean accuracy is also quantified, as shown in **Figure 3.3.** for four cases. For most of the pairs, there is a statistically significant decrease in performance.

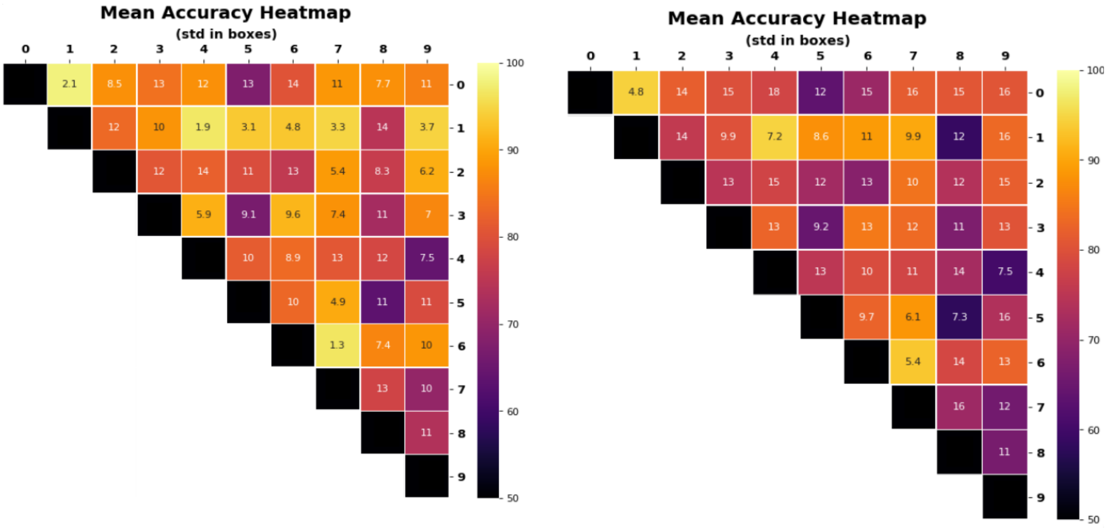


Figure 3.4. Global and local learning rate network performance. Mean accuracy heatmap. Every square represents one digit pair. The color of each box corresponds to the mean classification accuracy for the particular pair across all of its simulations and the number inside to standard deviation. Both networks do not utilize a stopping criterion and stop after 200 iterations.**(left)** Local learning rate network.**(right)** Global learning rate network.

3.4 Impact of Local Adaptive Learning Rate

To test the effect of the local adaptive learning rate, I compared the baseline network performance with the network's performance, while having a global learning rate equal to the maximum one, assuming that synapses grow with the same rate as they become larger. In both cases, the simulations were run without a stopping criterion since it depends on the learning rate's decrease. Thus, in both cases, the results are for 200 iterations of training so that they are comparable.

As in the previous section, the decrease in performance and increase in variance is noticeable by comparing the two heatmaps shown in **Figure 3.4.**. There is statistical significance for the difference in mean accuracy for some of the pairs, four of which are presented in **Figure 3.5.** as an example. The drop in performance is not as easily explained as in the previous section. However, a reasonable assumption could be that the local adaptive learning rate mechanism also makes the network less prone to errors due to initialization, but in a different way. In this model, synapses that have become large enough will most likely keep growing given the synaptic plasticity rule used. If the growth rate is constant, the initial large synapses will become too large fast, overshadowing smaller synapses and thus, limiting the network's capacity and hindering its performance.

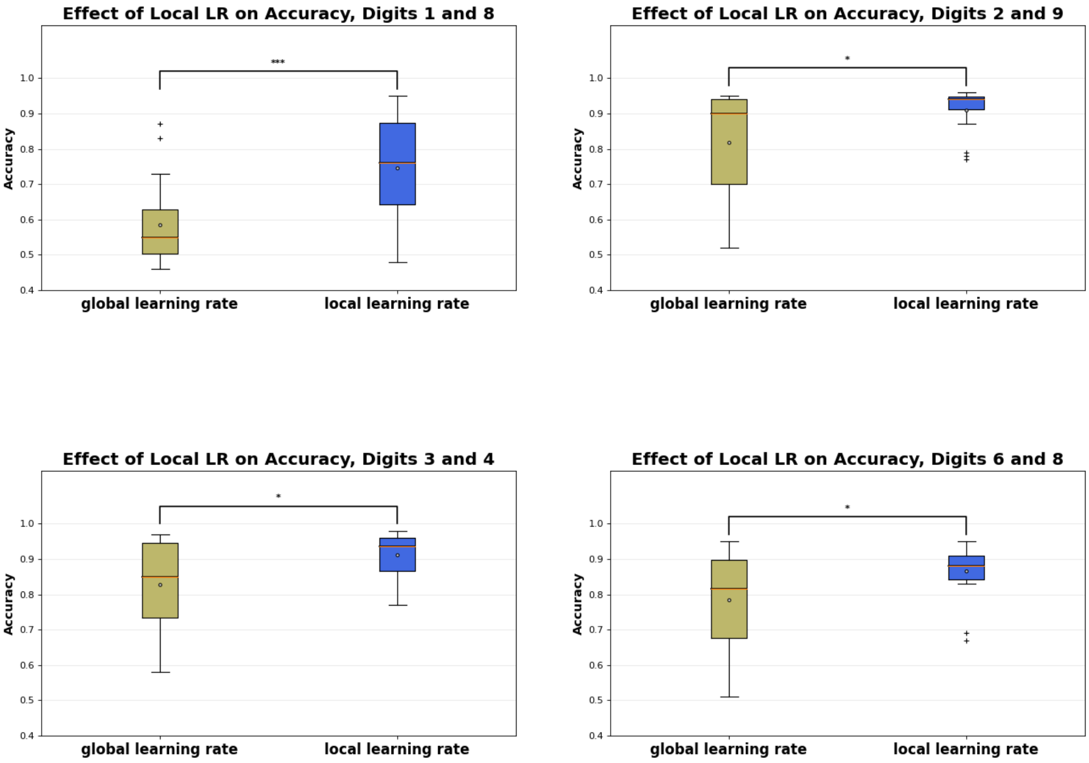


Figure 3.5. Comparison of global and local learning rate network performance. Each boxplot represents the accuracy across 18 simulations for the particular digit pair in the given condition. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, Welch t-test.

3.5 Impact of the Stopping Criterion

The baseline performance needs to be compared with that of an undertrained network and that of an overtrained one to evaluate the stopping criterion. Since the model runs for 100 to 200 iterations, with the current parameters and the stopping criterion, the undertrained network is trained for 50 iterations and the overtrained one for 550. The overtrained one is trained for over double the average baseline iterations because the local adaptive learning rate makes the model robust to overtraining since it slows down synaptic growth and a lot of iterations are needed to show its negative effects.

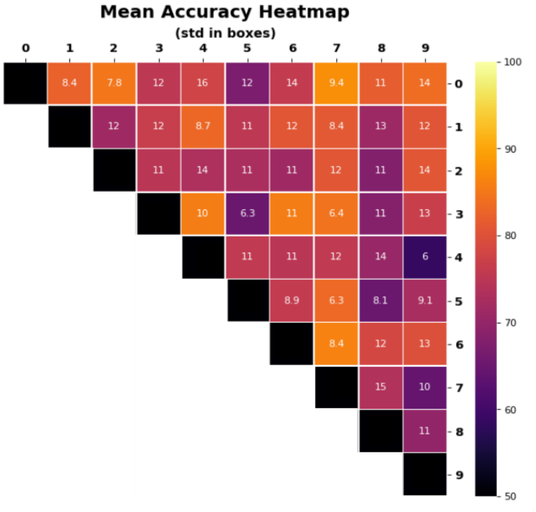


Figure 3.6. Undertrained and overtrained network performance. Mean accuracy heatmap. Every square represents one digit pair. The color of each box corresponds to the mean classification accuracy for the particular pair across all of its simulations and the number inside to standard deviation. Both networks do not utilize a stopping criterion, **(left)** the undertrained one stops after 50 iterations, and **(right)** the overtrained one stops after 550 iterations.

Again, heatmaps are shown for the undertrained and the overtrained networks in **Figure 3.6.** and example comparative boxplots for them and the baseline performance network in **Figure 3.7.**. Both of the examined cases lead to a drop in accuracy and an increase in variance but not in the same way.

The undertrained network does not have enough simulation time, leading to a drop in performance for almost every digit pair. With less simulation time, turnover happens less and that basically translates to network capacity decrease as stated above. Besides that, synapses don't have enough time to grow, and that way, memories are not solidified in the network, potentially leading to less accurate representations of a digit.

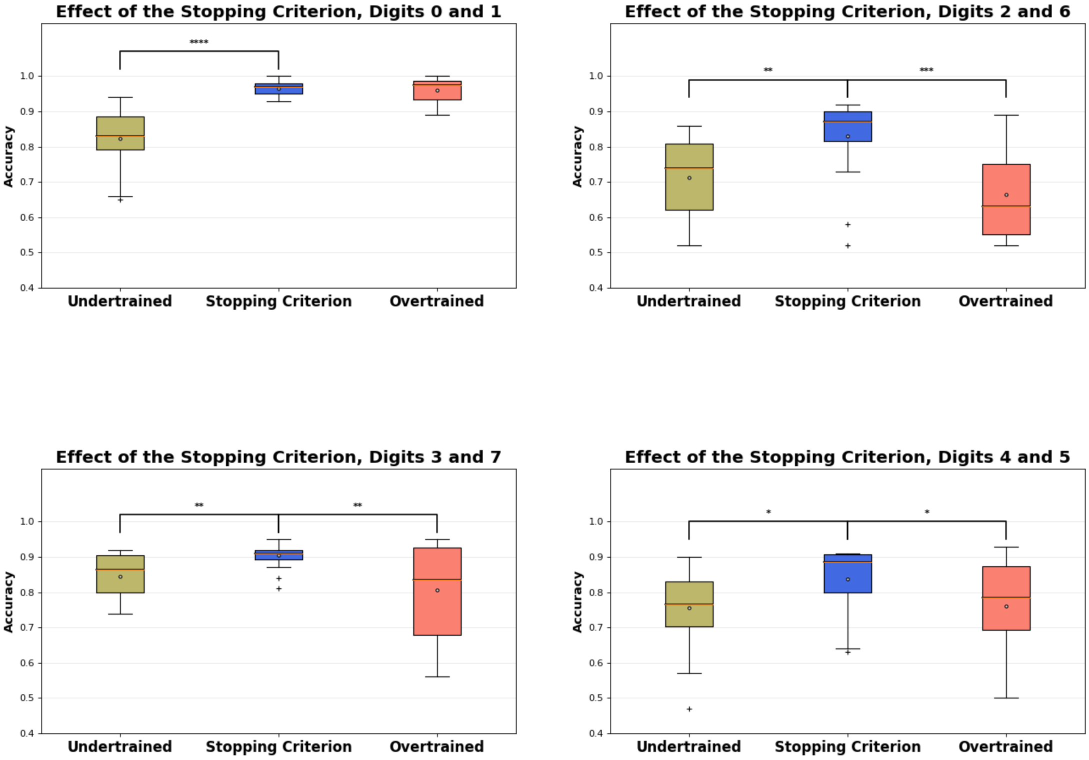


Figure 3.7. Comparison of undertrained and overtrained network performance to baseline. Each boxplot represents the accuracy across 18 simulations for the particular digit pair in the given condition. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, Welch t-test.

4 Discussion

Machine learning and neuroscience are two major scientific fields growing rapidly throughout the years, many times by feeding off of each other[35,155–159]. This work introduces a spiking neural network model that lies in the intersection between those two fields. It can classify handwritten digits in a binary framework while maintaining a phenomenologically modeled biophysical structure and utilizing multiple brain-inspired functionalities. Moreover, it can be characterized as a biologically plausible model due to numerous constraints incorporated in it that are present in the mammalian brain, like sparsity, firing frequency below 60Hz, sequential processing of information across time and a local learning rule. Given those restrictions, it is a model that can be used to test hypotheses regarding learning in the brain. This is accurately demonstrated by evaluating the contribution of three of its brain-inspired functionalities to learning and classification separately.

The neuro-inspired features that were evaluated in this study are the synaptic turnover, the local adaptive learning rate, and the bioinspired stopping criterion. All of them positively affect the network's classification performance as in their absence, accuracy drops, and variance increases. A plethora of neurobiological studies have already associated both synaptic turnover[137,138,141,142,144] and synaptic spine size and stability[147,149,150] with learning and memory. As far as synaptic spine stability goes, which corresponds to the local adaptive learning rate, quite a few deep learning algorithms have incorporated similar strategies with great success, the most popular one being Adam optimization algorithm[160].

As stated above, even when the proposed model utilizes the aforementioned mechanisms, it still does not perform as well or consistently as a standard deep learning algorithm. However, that was not the goal of this study. The network consists of a colossal hyperparameter space, which could potentially hide better parameter combinations. We did not search for those extensively since they were not required to show the advantages of the proposed neuronal mechanisms.

This network's advantage lies in efficiency. Even in this case, though, to showcase the suggested efficiency advantage, the network should be implemented in neuromorphic hardware. In standard high computing performance hardware, this model still falls behind machine learning algorithms, as the biological constraints hold it back. For example, a biological network with the aspect of time in it cannot process inputs in batches since each input needs to be processed serially.

What makes this model exciting and unique is that it provides an excellent template for evaluating how learning and memory work in the brain. One example of how this model could be utilized in that manner in the future is assessing its ability to learn continually. Continual learning is one of the elusive modern challenges in deep learning.

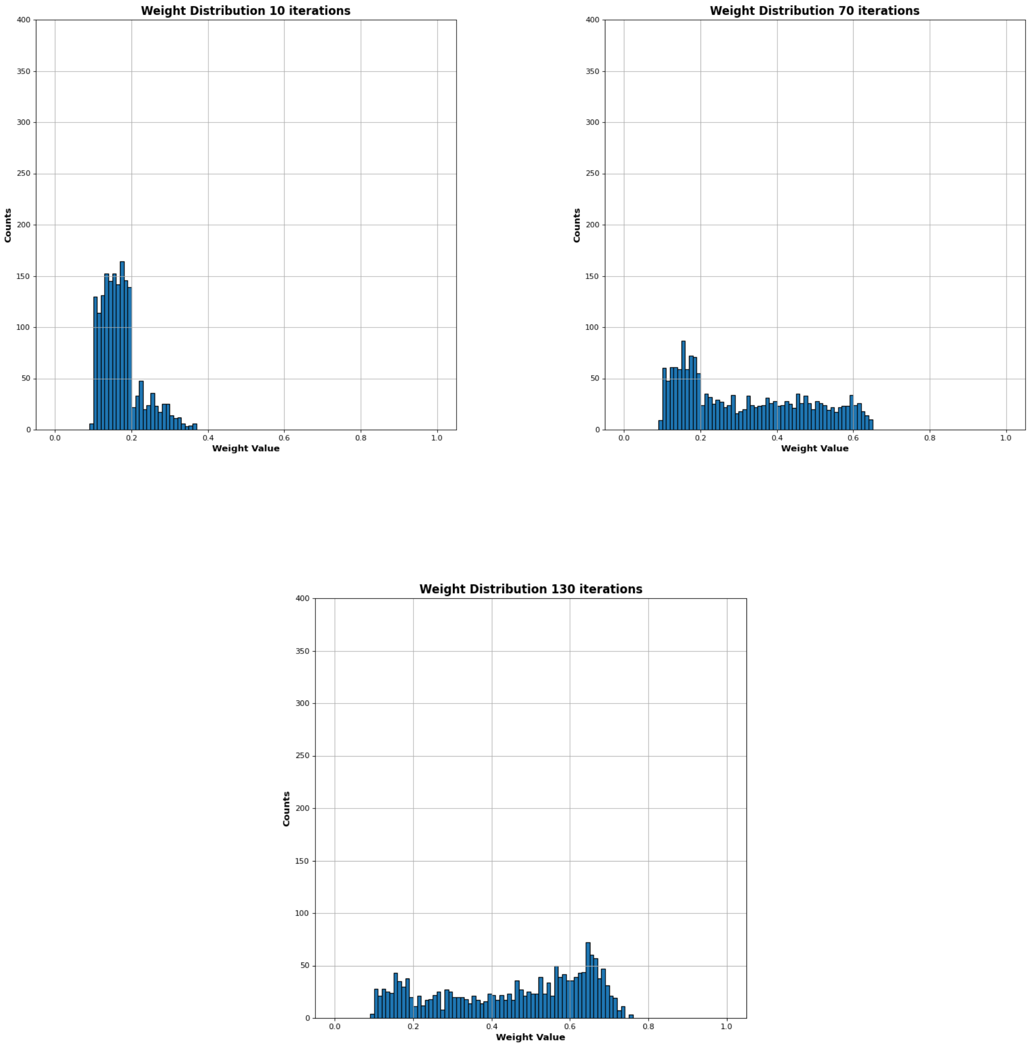
Continual or lifelong learning is defined as the ability to continually learn over time by accommodating new knowledge while retaining previously learned experiences[161]. Even though humans excel at learning continually and transferring knowledge from one task to another[162–164], artificial intelligence deep learning algorithms struggle with this problem for years[162,165,166]. The main issue that emerges in such algorithms is known as catastrophic forgetting. Learning new information leads to partially or, in the worst-case scenario, totally overwriting previously acquired knowledge[161,167].

One of the reasons this phenomenon occurs in deep neural networks is that gradient descent by backprop drives the network to a global minimum in the loss landscape, utilizing every potential resource of the network, regardless of its size. That said, when new information is added to the network, to learn it, it is forced to use computational nodes/neurons that already contain information, leading the network to overwrite it and, thus, forgetting. Many solutions have been proposed to this problem, one of the latest being the computation of a Fisher Information Matrix in a trained network to reveal which weights are the most important ones. Knowing that, those weights are frozen and the network is then trained on a new task with the rest of the available weights[168]. Even though this solution works, it is biologically implausible.

The proposed model could potentially solve catastrophic forgetting due to the lack of globality. Since it does not learn a particular task by utilizing every available resource, but by associating a subpopulation to a class with a local synaptic tag and capture plasticity rule, teaching it another class should be possible in theory if free synapses and subpopulations are available.

In conclusion, we present a biologically constrained model that utilizes biologically plausible mechanisms to perform image classification in a binary framework. Even though its performance is subpar compared to standard deep learning algorithms, its goal is not to outperform them but to highlight how certain neural functionalities contribute to learning and memory. Moreover, it should have efficiency advantages if implemented in neuromorphic hardware, and due to the locality of its learning rules, it could potentially solve the elusive problem of catastrophic forgetting.

A Supporting Figures for Results Chapter



Weight Distribution 10 iterations

Weight Distribution 70 iterations

Weight Distribution 130 iterations

Figure A.1. Weight distribution of a baseline network. Weight distribution after **(top left)**10 iterations, **(top right)**70 iterations and**(bottom)**130 iterations.

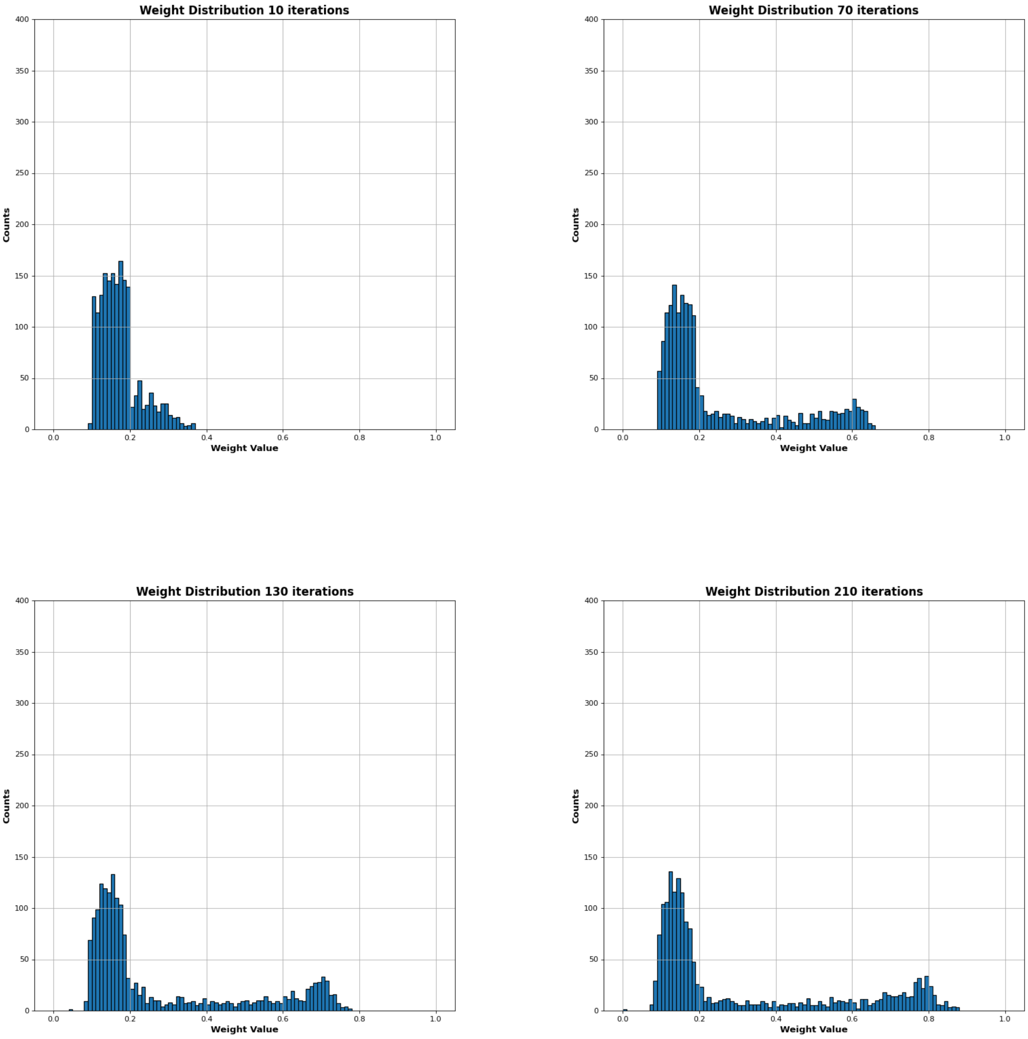


Figure A.2. Weight distribution of a no turnover network. Weight distribution after **(top left)**10 iterations, **(top right)**70 iterations, **(bottom left)**130 iterations and **(bottom right)**210 iterations. Clearly, some synapses remain unused, limiting the capacitance of the network

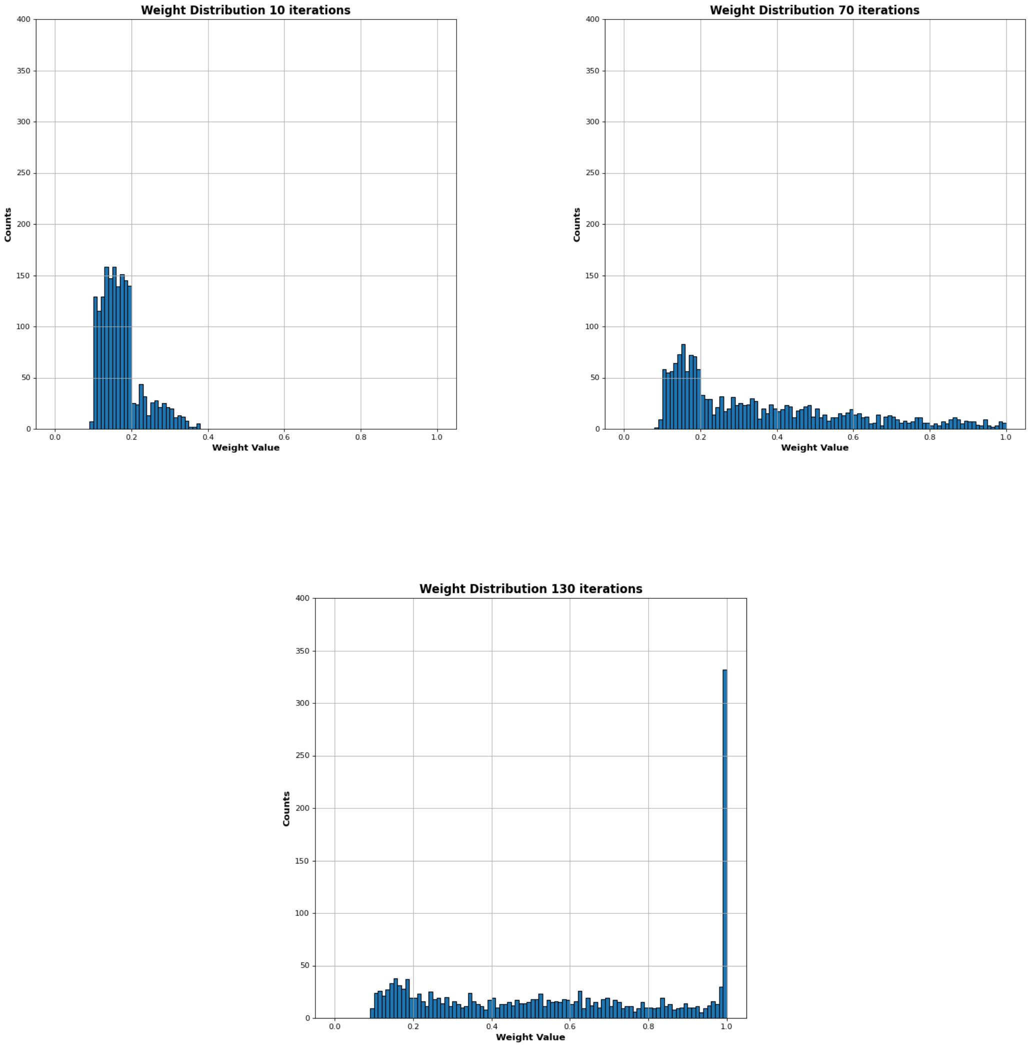
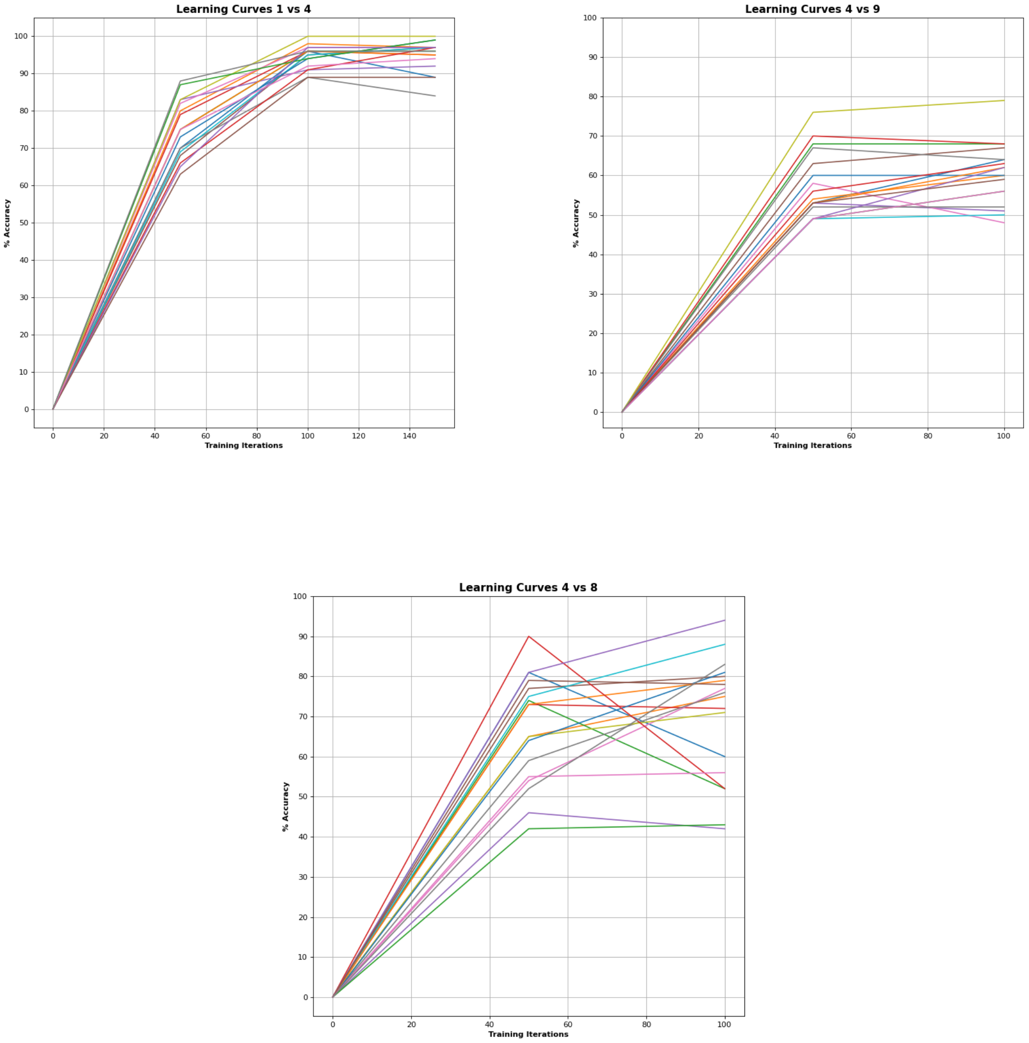


Figure A.3. Weight distribution of a global learning rate network. Weight distribution after **(top left)**10 iterations, **(top right)**70 iterations and**(bottom)**130 iterations. With a global learning rate, initially potentiated synapses grow too big too fast and overshadow the smaller ones.



Learning Curves 1 vs 4

Learning Curves 4 vs 9

Learning Curves 4 vs 8

Figure A.4. Featured examples of sample test Learning Curves. Featured examples of**(top left)** a pair that the network performs well on,**(top right)** a pair that the network doesn't perform well on and**(bottom)**a pair that the network performs inconsistently on.

List of Figures

References

1. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521, 436–444 (2015).
2. Lodish, H. F., Berk, A. & Kaiser, C. A. Molecular Cell Biology. (W.H. Freeman, 2012).
3. Bekkers, J. M. Pyramidal neurons. Curr. Biol. 21, R975 (2011).
4. Spruston, N. Pyramidal neurons: dendritic structure and synaptic integration. Nat. Rev. Neurosci. 9, 206–221 (2008).
5. Nelson, S. B., Hempel, C. & Sugino, K. Probing the transcriptome of neuronal cell types. Curr. Opin. Neurobiol. 16, 571–576 (2006).
6. Rudy, B., Fishell, G., Lee, S. & Hjerling-Leffler, J. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. Dev. Neurobiol. 71, 45–61 (2011).
7. Markram, H. et al. Interneurons of the neocortical inhibitory system. Nat. Rev. Neurosci. 5, 793–807 (2004).
8. Rudy, B. & McBain, C. J. Kv3 channels: voltage-gated K channels designed for high-frequency repetitive firing. Trends in Neurosciences vol. 24 517–526 (2001).
9. Tzilivaki, A., Kastellakis, G. & Poirazi, P. Challenging the point neuron dogma: FS basket cells as 2-stage nonlinear integrators. Nat. Commun. 10, 3664 (2019).
10. Ma, W.-P. et al. Visual representations by cortical somatostatin inhibitory neurons–selective but with weak and delayed responses. J. Neurosci. 30, 14371–14379 (2010).
11. Yavorska, I. & Wehr, M. Somatostatin-Expressing Inhibitory Interneurons in Cortical Circuits. Frontiers in Neural Circuits vol. 10 (2016).
12. Foster, M. A textbook of physiology / by M. Foster. (1888) doi:10.5962/bhl.title.1854.
13. Silverthorn, D. U. Human Physiology: An Integrated Approach. (Benjamin-Cummings Publishing Company, 2010).
14. Südhof, T. C. Towards an Understanding of Synapse Formation. Neuron 100, 276–293 (2018).

15. Krugers, H. Regulation of excitatory synapses and fearful memories by stress hormones. Frontiers in Behavioral Neuroscience vol. 5 (2011).

16. Bosman, C. A. & Aboitiz, F. Functional constraints in the evolution of brain circuits. Front. Neurosci. 9, 303 (2015).

17. Poirazi, P., Brannon, T. & Mel, B. W. Pyramidal neuron as two-layer neural network. Neuron 37, 989–999 (2003).

18. Poirazi, P. & Mel, B. W. Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. Neuron 29, 779–796 (2001).

19. Kastellakis, G., Silva, A. J. & Poirazi, P. Linking Memories across Time via Neuronal and Dendritic Overlaps in Model Neurons with Active Dendrites. Cell Rep. 17, 1491–1504 (2016).

20. Larkum, M. E., Zhu, J. J. & Sakmann, B. A new cellular mechanism for coupling inputs arriving at different cortical layers. Nature 398, 338–341 (1999).

21. Gidon, A. et al. Dendritic action potentials and computation in human layer 2/3 cortical neurons. Science 367, 83–87 (2020).

22. Vervaeke, K., Lorincz, A., Nusser, Z. & Silver, R. A. Gap Junctions Compensate for Sublinear Dendritic Integration in an Inhibitory Network. Science vol. 335 1624–1628 (2012).

23. Abrahamsson, T., Cathala, L., Matsui, K., Shigemoto, R. & DiGregorio, D. A. Thin Dendrites of Cerebellar Interneurons Confer Sublinear Synaptic Integration and a Gradient of Short-Term Plasticity. Neuron vol. 73 1159–1172 (2012).

24. Cash, S. & Yuste, R. Linear summation of excitatory inputs by CA1 pyramidal neurons. Neuron 22, 383–394 (1999).

25. Hu, H., Martina, M. & Jonas, P. Dendritic mechanisms underlying rapid synaptic activation of fast-spiking hippocampal interneurons. Science 327, 52–58 (2010).

26. Johnston, D. & Narayanan, R. Active dendrites: colorful wings of the mysterious butterflies. Trends in Neurosciences vol. 31 309–316 (2008).

27. Agmon-Snir, H., Carr, C. E. & Rinzel, J. The role of dendrites in auditory coincidence detection. Nature 393, 268–272 (1998).

28. Mel, B. W., Ruderman, D. L. & Archie, K. A. Translation-Invariant Orientation Tuning in Visual ‘Complex’ Cells Could Derive from Intradendritic Computations. The Journal of Neuroscience vol. 18 4325–4334 (1998).

29. Archie, K. A. & Mel, B. W. A model for intradendritic computation of binocular disparity. Nature Neuroscience vol. 3 54–63 (2000).

30. Tran-Van-Minh, A. et al. Contribution of sublinear and supralinear dendritic integration to neuronal computations. Front. Cell. Neurosci. 9, 67 (2015).

31. Lee, D., Lin, B.-J. & Lee, A. K. Hippocampal place fields emerge upon single-cell manipulation of excitability during behavior. Science 337, 849–853 (2012).

32. Xu, N.-L. et al. Nonlinear dendritic integration of sensory and motor input during an active sensing task. Nature 492, 247–251 (2012).

33. Lavzin, M., Rapoport, S., Polsky, A., Garion, L. & Schiller, J. Nonlinear dendritic processing determines angular tuning of barrel cortex neurons in vivo. Nature 490, 397–401 (2012).

34. Longordo, F., To, M.-S., Ikeda, K. & Stuart, G. J. Sublinear integration underlies binocular processing in primary visual cortex. Nat. Neurosci. 16, 714–723 (2013).

35. Guerguiev, J., Lillicrap, T. P. & Richards, B. A. Towards deep learning with segregated dendrites. Elife 6, (2017).

36. Payeur, A., Guerguiev, J., Zenke, F., Richards, B. A. & Naud, R. Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. Nat. Neurosci. 24, 1010–1019 (2021).

37. Sacramento, J., Costa, R. P., Bengio, Y. & Senn, W. Dendritic error backpropagation in deep cortical microcircuits. arXiv (2018).

38. Tonegawa, S., Pignatelli, M., Roy, D. S. & Ryan, T. J. Memory engram storage and retrieval. Curr. Opin. Neurobiol. 35, 101–109 (2015).

39. Josselyn, S. A., Köhler, S. & Frankland, P. W. Finding the engram. Nature Reviews Neuroscience vol. 16 521–534 (2015).

40. Ryan, T. J., Roy, D. S., Pignatelli, M., Arons, A. & Tonegawa, S. Memory. Engram cells retain memory under retrograde amnesia. Science 348, 1007–1013 (2015).

41. Liu, X. et al. Optogenetic stimulation of a hippocampal engram activates fear memory recall. Nature vol. 484 381–385 (2012).

42. Roy, D. S. et al. Memory retrieval by activating engram cells in mouse models of early Alzheimer's disease. Nature vol. 531 508–512 (2016).

43. Semon, R. THE MNEME. The Journal of Nervous and Mental Disease vol. 62 332 (1925).

44. Lashley, K. S. 1. In search of the engram. Brain Physiology and Psychology 1–32 (1966) doi:10.1525/9780520318267-001.

45. Penfield, W. & Rasmussen, T. The Cerebral Cortex of Man: A Clinical Study of Localization of Function. (1955).

46. Scoville, W. B. & Milner, B. Loss of recent memory after bilateral hippocampal lesions. J. Neurol. Neurosurg. Psychiatry 20, 11–21 (1957).

47. Morris, R. G. M., Garrud, P., Rawlins, J. N. P. & O'Keefe, J. Place navigation impaired in rats with hippocampal lesions. Nature vol. 297 681–683 (1982).

48. Moser, M.-B. & Moser, E. I. Distributed Encoding and Retrieval of Spatial Memory in the Hippocampus. The Journal of Neuroscience vol. 18 7535–7542 (1998).

49. The Croonian lecture.—La fine structure des centres nerveux. Proceedings of the Royal Society of London vol. 55 444–468 (1894).

50. Attneave, F., B., M. & Hebb, D. O. The Organization of Behavior; A Neuropsychological Theory. The American Journal of Psychology vol. 63 633 (1950).

51. Masquelier, T., Guyonneau, R. & Thorpe, S. J. Competitive STDP-based spike pattern learning. Neural Comput. 21, 1259–1276 (2009).

52. Clopath, C., Büsing, L., Vasilaki, E. & Gerstner, W. Connectivity reflects coding: a model of voltage-based STDP with homeostasis. Nature Neuroscience vol. 13 344–352 (2010).

53. Iakymchuk, T., Rosado-Muñoz, A., Guerrero-Martínez, J. F., Bataller-Mompeán, M. & Francés-Víllora, J. V. Simplified spiking neural network architecture and STDP learning algorithm applied to image classification. EURASIP Journal on Image and Video Processing vol. 2015 (2015).

54. Krunglevicius, D. STDP Learning Under Variable Noise Levels. Proceedings of the International Conference on Neural Computation Theory and Applications (2014) doi:10.5220/0005072401650171.

55. Abigail, M. Continuous variable models of voltage-based STDP. Frontiers in Computational Neuroscience vol. 6 (2012).

56. Ebong, I. & Mazumder, P. Memristor based STDP learning network for position detection. 2010 International Conference on Microelectronics (2010) doi:10.1109/icm.2010.5696142.

57. Bi, G. Q. & Poo, M. M. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. J. Neurosci. 18, 10464–10472 (1998).

58. Frey, U. & Morris, R. G. Synaptic tagging and long-term potentiation. Nature 385, 533–536 (1997).

59. Redondo, R. L. & Morris, R. G. M. Making memories last: the synaptic tagging and capture hypothesis. Nat. Rev. Neurosci. 12, 17–30 (2011).

60. Banko, J. L. & Klann, E. Cap-dependent translation initiation and memory. Prog. Brain Res. 169, 59–80 (2008).

61. Steward, O., Wallace, C. S., Lyford, G. L. & Worley, P. F. Synaptic activation causes the mRNA for the IEG Arc to localize selectively near activated postsynaptic sites on dendrites. Neuron 21, 741–751 (1998).

62. Michmizos, D., Koutsouraki, E., Asprodini, E. & Baloyannis, S. Synaptic plasticity: a unifying model to address some persisting questions. Int. J. Neurosci. 121, 289–304 (2011).

63. Andersen, N., Krauth, N. & Nabavi, S. Hebbian plasticity in vivo: relevance and induction. Curr. Opin. Neurobiol. 45, 188–192 (2017).

64. Bernardinelli, Y., Nikonenko, I. & Muller, D. Structural plasticity: mechanisms and contribution to developmental psychiatric disorders. Front. Neuroanat. 8, 123 (2014).

65. La Rosa, C., Parolisi, R. & Bonfanti, L. Brain Structural Plasticity: From Adult Neurogenesis to Immature Neurons. Front. Neurosci. 14, 75 (2020).

66. Raisman, G. Neuronal plasticity in the septal nuclei of the adult rat. Brain Res. 14, 25–48 (1969).

67. Jones, T. A., Chu, C. J., Grande, L. A. & Gregory, A. D. Motor skills training enhances lesion-induced structural plasticity in the motor cortex of adult rats. J. Neurosci. 19, 10153–10163 (1999).

68. Butz, M., Wörgötter, F. & van Ooyen, A. Activity-dependent structural plasticity. Brain Research Reviews vol. 60 287–305 (2009).

69. Turrigiano, G., Abbott, L. F. & Marder, E. Activity-dependent changes in the intrinsic properties of cultured neurons. Science 264, 974–977 (1994).

70. MacLean, J. N. et al. Activity-Independent Coregulation of IA and Ih in Rhythmically Active Neurons. Journal of Neurophysiology vol. 94 3601–3617 (2005).

71. Temporal, S., Lett, K. M. & Schulz, D. J. Activity-dependent feedback regulates correlated ion channel mRNA levels in single identified motor neurons. Curr. Biol. 24, 1899–1904 (2014).

72. Hudson, A. E., Gollnick, C., Gourdine, J.-P. & Prinz, A. A. Degradation of extracellular chondroitin sulfate delays recovery of network activity after perturbation. J. Neurophysiol. 114, 1346–1352 (2015).

73. Northcutt, A. J. & Schulz, D. J. Molecular mechanisms of homeostatic plasticity in central pattern generator networks. Dev. Neurobiol. 80, 58–69 (2020).

74. Turrigiano, G. G. & Nelson, S. B. Homeostatic plasticity in the developing nervous system. Nature Reviews Neuroscience vol. 5 97–107 (2004).

75. Mitchell, T. M. Machine Learning. (1997).

76. Russell, S. J., Russell, S. J., Norvig, P. & Davis, E. Artificial Intelligence: A Modern Approach. (Prentice Hall, 2010).

77. Cortes, C. & Vapnik, V. Support-vector networks. Machine Learning vol. 20 273–297 (1995).

78. Altman, N. S. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. The American Statistician vol. 46 175 (1992).

79. Wu, X. & Kumar, V. The Top Ten Algorithms in Data Mining. (CRC Press, 2009).

80. Hinton, G. E. & Sejnowski, T. J. Unsupervised Learning: Foundations of Neural Computation. (MIT Press, 1999).

81. Hinton, G. E. A Practical Guide to Training Restricted Boltzmann Machines. Lecture Notes in Computer Science 599–619 (2012) doi:10.1007/978-3-642-35289-8_32.

82. Pearson, K. On Lines and Planes of Closest Fit to Systems of Points in Space. (1901).

83. Lloyd, S. Least squares quantization in PCM. IEEE Transactions on Information Theory vol. 28 129–137 (1982).

84. Schmidhuber, J. Deep learning in neural networks: an overview. Neural Netw. 61, 85–117 (2015).

85. Bengio, Y., Courville, A. & Vincent, P. Representation learning: a review and new perspectives. IEEE Trans. Pattern Anal. Mach. Intell. 35, 1798–1828 (2013).

86. McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biology vol. 52 99–115 (1990).

87. Haykin, S. S. Neural Networks and Learning Machines. (Prentice Hall, 2009).

88. Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol. Rev. 65, 386–408 (1958).

89. LeCun, Y. & Bengio, Y. Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks (1995).

90. Fukushima, K. Neocognitron. Scholarpedia vol. 2 1717 (2007).

91. Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat's striate cortex. The Journal of Physiology vol. 148 574–591 (1959).

92. LeCun, Y. et al. Backpropagation Applied to Handwritten Zip Code Recognition. Neural Computation vol. 1 541–551 (1989).

93. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J. Mach. Learn. Res. (2014).

94. Vaswani, A. et al. Attention Is All You Need. arXiv.

95. Colman, A. M. A Dictionary of Psychology. (Oxford University Press, USA, 2015).

96. Hunsberger, E. Spiking Deep Neural Networks: Engineered and Biological Approaches to Object Recognition. (2017).

97. Lillicrap, T. P., Cownden, D., Tweed, D. B. & Akerman, C. J. Random synaptic feedback weights support error backpropagation for deep learning. Nat. Commun. 7, 13276 (2016).

98. Maass, W. Networks of spiking neurons: The third generation of neural network models. Neural Networks vol. 10 1659–1671 (1997).

99. Virgilio G, C. D., Sossa A, J. H., Antelis, J. M. & Falcón, L. E. Spiking Neural Networks applied to the classification of motor tasks in EEG signals. Neural Netw. 122, 130–143 (2020).

100. Hodgkin, A. L. & Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. 1952. Bull. Math. Biol. 52, 25–71; discussion 5–23 (1990).

101. Abbott, L. F. Lapicque's introduction of the integrate-and-fire model neuron (1907). Brain Res. Bull. 50, 303–304 (1999).

102. Izhikevich, E. M. Simple model of spiking neurons. IEEE Trans. Neural Netw. 14, 1569–1572 (2003).

103. Gerstner, W. & Kistler, W. M. Spiking Neuron Models: Single Neurons, Populations, Plasticity. (Cambridge University Press, 2002).

104. Gerstner, W., Kistler, W. M., Naud, R. & Paninski, L. Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition. (Cambridge University Press, 2014).

105. Richardson, M. J. E., Brunel, N. & Hakim, V. From Subthreshold to Firing-Rate Resonance. Journal of Neurophysiology vol. 89 2538–2554 (2003).

106. Fourcaud-Trocmé, N., Hansel, D., van Vreeswijk, C. & Brunel, N. How spike generation mechanisms determine the neuronal response to fluctuating inputs. J. Neurosci. 23, 11628–11640 (2003).

107. Brette, R. & Gerstner, W. Adaptive Exponential Integrate-and-Fire Model as an Effective Description of Neuronal Activity. Journal of Neurophysiology vol. 94 3637–3642 (2005).

108. Zenke, F., Agnes, E. J. & Gerstner, W. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. Nat. Commun. 6, 6922 (2015).

109. Diehl, P. U. & Cook, M. Unsupervised learning of digit recognition using spike-timing-dependent plasticity. Front. Comput. Neurosci. 9, 99 (2015).

110. Kheradpisheh, S. R., Ganjtabesh, M., Thorpe, S. J. & Masquelier, T. STDP-based spiking deep convolutional neural networks for object recognition. Neural Netw. 99, 56–67 (2018).

111. Neftci, E. O., Mostafa, H. & Zenke, F. Surrogate Gradient Learning in Spiking Neural Networks: Bringing the Power of Gradient-Based Optimization to Spiking Neural Networks. IEEE Signal Processing Magazine vol. 36 51–63 (2019).

112. Jolivet, R., Rauch, A., Lüscher, H.-R. & Gerstner, W. Predicting spike timing of neocortical pyramidal neurons by simple threshold models. J. Comput. Neurosci. 21, 35–49 (2006).

113. Higley, M. J. & Sabatini, B. L. Calcium signaling in dendritic spines. Cold Spring Harb. Perspect. Biol. 4, a005686 (2012).

114. Kang, H. & Schuman, E. M. A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity. Science 273, 1402–1406 (1996).

115. Huber, K. M., Kayser, M. S. & Bear, M. F. Role for rapid dendritic protein synthesis in hippocampal mGluR-dependent long-term depression. Science 288, 1254–1257 (2000).

116. Frey, U., Krug, M., Brödemann, R., Reymann, K. & Matthies, H. Long-term potentiation induced in dendrites separated from rat's CA1 pyramidal somata does not establish a late phase. Neurosci. Lett. 97, 135–139 (1989).

117. Nguyen, P. V., Abel, T. & Kandel, E. R. Requirement of a critical period of transcription for induction of a late phase of LTP. Science 265, 1104–1107 (1994).

118. Morris, R. G. & Frey, U. Hippocampal synaptic plasticity: role in spatial learning or the automatic recording of attended experience? Philos. Trans. R. Soc. Lond. B Biol. Sci. 352, 1489–1503 (1997).

119. Shouval, H. Z., Bear, M. F. & Cooper, L. N. A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. Proc. Natl. Acad. Sci. U. S. A. 99, 10831–10836 (2002).

120. Disterhoft, J. F. & Oh, M. M. Learning, aging and intrinsic neuronal plasticity. Trends Neurosci. 29, 587–599 (2006).

121. Silva, A. J., Zhou, Y., Rogerson, T., Shobe, J. & Balaji, J. Molecular and cellular approaches to memory allocation in neural circuits. Science 326, 391–395 (2009).

122. Zhou, Y. et al. CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. Nat. Neurosci. 12, 1438–1443 (2009).

123. Frick, A., Magee, J. & Johnston, D. LTP is accompanied by an enhanced local excitability of pyramidal neuron dendrites. Nat. Neurosci. 7, 126–135 (2004).

124. Oh, M. M., Kuo, A. G., Wu, W. W., Sametsky, E. A. & Disterhoft, J. F. Watermaze learning enhances excitability of CA1 pyramidal neurons. J. Neurophysiol. 90, 2171–2179 (2003).

125. Sehgal, M., Ehlers, V. L. & Moyer, J. R., Jr. Learning enhances intrinsic excitability in a subset of lateral amygdala neurons. Learn. Mem. 21, 161–170 (2014).

126. Huang, Y. H. et al. CREB modulates the functional output of nucleus accumbens neurons: a critical role of N-methyl-D-aspartate glutamate receptor (NMDAR) receptors. J. Biol. Chem. 283, 2751–2760 (2008).

127. Kim, J., Kwon, J.-T., Kim, H.-S. & Han, J.-H. CREB and neuronal selection for memory trace. Front. Neural Circuits 7, 44 (2013).

128. Dong, Y. et al. CREB modulates excitability of nucleus accumbens neurons. Nat. Neurosci. 9, 475–477 (2006).

129. Han, M.-H. et al. Role of cAMP response element-binding protein in the rat locus ceruleus: regulation of neuronal activity and opiate withdrawal behaviors. J. Neurosci. 26, 4624–4629 (2006).

130. Lopez de Armentia, M. et al. cAMP response element-binding protein-mediated gene expression increases the intrinsic excitability of CA1 pyramidal neurons. J. Neurosci. 27, 13909–13918 (2007).

131. Silva, A. J., Kogan, J. H., Frankland, P. W. & Kida, S. CREB and memory. Annu. Rev. Neurosci. 21, 127–148 (1998).

132. Benito, E. & Barco, A. CREB's control of intrinsic and synaptic plasticity: implications for CREB-dependent memory models. Trends Neurosci. 33, 230–240 (2010).

133. Kim, J., Kwon, J.-T., Kim, H.-S., Josselyn, S. A. & Han, J.-H. Memory recall and modifications by activating neurons with elevated CREB. Nat. Neurosci. 17, 65–72 (2014).

134. Rogerson, T. et al. Synaptic tagging during memory allocation. Nat. Rev. Neurosci. 15, 157–169 (2014).

135. Turrigiano, G. G. The self-tuning neuron: synaptic scaling of excitatory synapses. Cell 135, 422–435 (2008).

136. Deng, L. The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web]. IEEE Signal Processing Magazine vol. 29 141–142 (2012).

137. Frank, A. C. et al. Hotspots of dendritic spine turnover facilitate clustered spine addition and learning and memory. Nat. Commun. 9, 422 (2018).

138. Trachtenberg, J. T. et al. Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. Nature 420, 788–794 (2002).

139. Kastellakis, G. & Poirazi, P. Synaptic Clustering and Memory Formation. Front. Mol. Neurosci. 12, 300 (2019).

140. Schmidt, S. et al. Experience-dependent structural plasticity in the adult brain: How the learning brain grows. Neuroimage 225, 117502 (2021).

141. Holtmaat, A. J. G. D. et al. Transient and persistent dendritic spines in the neocortex in vivo. Neuron 45, 279–291 (2005).

142. Holtmaat, A. & Svoboda, K. Experience-dependent structural synaptic plasticity in the mammalian brain. Nat. Rev. Neurosci. 10, 647–658 (2009).

143. Fu, M., Yu, X., Lu, J. & Zuo, Y. Repetitive motor learning induces coordinated formation of clustered dendritic spines in vivo. Nature 483, 92–95 (2012).

144. Moczulska, K. E. et al. Dynamics of dendritic spines in the mouse auditory cortex during memory formation and memory recall. Proc. Natl. Acad. Sci. U. S. A. 110, 18315–18320 (2013).

145. Yang, G., Pan, F. & Gan, W.-B. Stably maintained dendritic spines are associated with lifelong memories. Nature 462, 920–924 (2009).

146. Stepanyants, A., Hof, P. R. & Chklovskii, D. B. Geometry and structural plasticity of synaptic connectivity. Neuron 34, 275–288 (2002).

147. Matsuzaki, M., Honkura, N., Ellis-Davies, G. C. R. & Kasai, H. Structural basis of long-term potentiation in single dendritic spines. Nature 429, 761–766 (2004).

148. Matus, A. Growth of dendritic spines: a continuing story. Curr. Opin. Neurobiol. 15, 67–72 (2005).

149. Kasai, H., Matsuzaki, M., Noguchi, J., Yasumatsu, N. & Nakahara, H. Structure-stability-function relationships of dendritic spines. Trends Neurosci. 26, 360–368 (2003).

150. Grutzendler, J., Kasthuri, N. & Gan, W.-B. Long-term dendritic spine stability in the adult cortex. Nature 420, 812–816 (2002).

151. Harris, C. R. et al. Array programming with NumPy. Nature 585, 357–362 (2020).

152. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17, 261–272 (2020).

153. Waskom, M. seaborn: statistical data visualization. Journal of Open Source Software vol. 6 3021 (2021).

154. Hunter, J. D. Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering vol. 9 90–95 (2007).

155. Richards, B. A. et al. A deep learning framework for neuroscience. Nat. Neurosci. 22, 1761–1770 (2019).

156. Savage, N. How AI and neuroscience drive each other forwards. Nature vol. 571 S15–S17 (2019).

157. Marblestone, A. H., Wayne, G. & Kording, K. P. Toward an Integration of Deep Learning and Neuroscience. Front. Comput. Neurosci. 10, 94 (2016).

158. Chavlis, S. & Poirazi, P. Drawing inspiration from biological dendrites to empower artificial neural networks. Curr. Opin. Neurobiol. 70, 1–10 (2021).

159. Mathis, A. et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat. Neurosci. 21, 1281–1289 (2018).

160. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. in Proceedings of the 3rd International Conference on Learning Representations (ICLR) (2014).

161. Parisi, G. I., Kemker, R., Part, J. L., Kanan, C. & Wermter, S. Continual lifelong learning with neural networks: A review. Neural Netw. 113, 54–71 (2019).

162. Hassabis, D., Kumaran, D., Summerfield, C. & Botvinick, M. Neuroscience-Inspired Artificial Intelligence. Neuron 95, 245–258 (2017).

163. Barnett, S. M. & Ceci, S. J. When and where do we apply what we learn? A taxonomy for far transfer. Psychol. Bull. 128, 612–637 (2002).

164. Holyoak, K. J. & Thagard, P. The analogical mind. Am. Psychol. 52, 35–44 (1997).

165. Thrun, S. & Mitchell, T. M. Lifelong Robot Learning. The Biology and Technology of Intelligent Autonomous Agents 165–196 (1995) doi:10.1007/978-3-642-79629-6_7.

166. Thrun, S. & Mitchell, T. M. Learning One More Thing. (1994) doi:10.21236/ada285342.

167. McCloskey, M. & Cohen, N. J. Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem. Psychology of Learning and Motivation 109–165 (1989) doi:10.1016/s0079-7421(08)60536-8.

168. Kirkpatrick, J. et al. Overcoming catastrophic forgetting in neural networks. Proc. Natl. Acad. Sci. U. S. A. 114, 3521–3526 (2017).