



Categorization and decision-making in a neurobiologically plausible spiking network using a STDP-like learning rule



Michael Beyeler^{a,*}, Nikil D. Dutt^a, Jeffrey L. Krichmar^{a,b}

^a Department of Computer Science, University of California Irvine, Irvine, CA 92697-3435, United States

^b Department of Cognitive Sciences, 3151 Social Science Plaza, University of California Irvine, Irvine, CA 92697-5100, United States

ARTICLE INFO

Article history:

Received 24 February 2013

Received in revised form 28 July 2013

Accepted 31 July 2013

Keywords:

Spiking neural network

Synaptic dynamics

STDP

Supervised learning

Object recognition

Decision-making

ABSTRACT

Understanding how the human brain is able to efficiently perceive and understand a visual scene is still a field of ongoing research. Although many studies have focused on the design and optimization of neural networks to solve visual recognition tasks, most of them either lack neurobiologically plausible learning rules or decision-making processes. Here we present a large-scale model of a hierarchical spiking neural network (SNN) that integrates a low-level memory encoding mechanism with a higher-level decision process to perform a visual classification task in real-time. The model consists of Izhikevich neurons and conductance-based synapses for realistic approximation of neuronal dynamics, a spike-timing-dependent plasticity (STDP) synaptic learning rule with additional synaptic dynamics for memory encoding, and an accumulator model for memory retrieval and categorization. The full network, which comprised 71,026 neurons and approximately 133 million synapses, ran in real-time on a single off-the-shelf graphics processing unit (GPU). The network was constructed on a publicly available SNN simulator that supports general-purpose neuromorphic computer chips. The network achieved 92% correct classifications on MNIST in 100 rounds of random sub-sampling, which is comparable to other SNN approaches and provides a conservative and reliable performance metric. Additionally, the model correctly predicted reaction times from psychophysical experiments. Because of the scalability of the approach and its neurobiological fidelity, the current model can be extended to an efficient neuromorphic implementation that supports more generalized object recognition and decision-making architectures found in the brain.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Object recognition in monkeys has traditionally been associated with an anatomically distinct pathway termed the “what” (or ventral) visual stream (Ungerleider & Haxby, 1994), which consists of at least V1, V2, V4, and various regions in the inferior and anterior temporal cortices (e.g., TEO, TE1–TE3, TEa, and TEM) (Rolls, 2012; Rolls & Deco, 2002). While traveling along this pathway, the characteristics of the stimuli to which neurons respond become more complex (Rolls, 2012; Ungerleider & Haxby, 1994), ranging from rather simple stimuli with small receptive fields such as oriented bars in V1 (Hubel & Wiesel, 1965) to relatively large and more abstract objects such as faces in the inferotemporal cortex (IT) (Bruce, Desimone, & Gross, 1981). These empirical observations have led to a number of classic studies modeling the

ventral stream as a hierarchical feed-forward network, such as the Neocognitron (Fukushima, 1980), HMAX (Riesenhuber & Poggio, 1999), or VisNet (Rolls, 2012; Wallis & Rolls, 1997)—although it should be noted that the notion of a strictly hierarchical or feed-forward network has been questioned by recent anatomical studies that reserve a more important functional role for bi-directional and non-hierarchical connections (Markov et al., 2012, 2011). Inspired by these classic models, a variety of more conventional machine learning algorithms have emerged that demonstrate the extraordinary performance in certain recognition tasks, such as convolutional neural networks (CNNs) in handwriting recognition (Ciresan, Meier, Masci, & Schmidhuber, 2011; LeCun, Bottou, Bengio, & Haffner, 1998; Simard, Steinkraus, & Platt, 2003)—or for that matter, adaptive boosting in face recognition (Viola & Jones, 2001). Although CNNs implement a network topology that is biologically-inspired, they often rely on the error backpropagation (gradient descent), which has been criticized for being biologically unrealistic because it involves variables that cannot be computed locally (Rolls & Deco, 2002). Part of the challenge is thus to discover how comparably hard problems can be solved by more biologically plausible networks relying on local learning rules that operate on the abstraction level of a synapse.

* Correspondence to: Department of Cognitive Sciences, 2236 Social & Behavioral Sciences Gateway, University of California Irvine, Irvine, CA 92697-5100, United States.

E-mail address: mbeyeler@uci.edu (M. Beyeler).

URL: <http://sites.uci.edu/mbeyeler> (M. Beyeler).

A potential candidate for such a mechanism is spike-timing-dependent plasticity (STDP), (Bi & Poo, 2001; Sjöström, Turrigiano, & Nelson, 2001; Song, Miller, & Abbott, 2000), a paradigm which modulates the weight of synapses according to their degree of causality. Many different variants of STDP seem to exist in the brain, and many different models to explain them have emerged over the years (Morrison, Diesmann, & Gerstner, 2008). In an effort to implement STDP-like learning rules using only information locally available at the synapse without algorithmically storing the spike timings, several models have proposed to pair presynaptic spiking with postsynaptic voltage, determining the weight change by using either the temporal change of postsynaptic voltage (Porr, Saudargiene, & Worgotter, 2004), a piece-wise linear function to approximate postsynaptic voltage (Gorchetchnikov, Versace, & Hasselmo, 2005) or postsynaptic Calcium concentration (Brader, Senn, & Fusi, 2007; Graupner & Brunel, 2012). Networks with STDP have been shown to be able to learn precise spike times through supervised learning (Legenstein, Naeger, & Maass, 2005; Pfister, Toyozumi, Barber, & Gerstner, 2006), to implement reinforcement learning (Florian, 2007; Izhikevich, 2007b; O'Brien & Srinivasa, 2013), to develop localized receptive fields (Clopath, Busing, Vasilaki, & Gerstner, 2010), or to classify highly correlated patterns of neuronal activity (Brader et al., 2007).

Once an internal representation of a visual object is built in the brain, the question then remains how this memory can be retrieved from the system in order to make a perceptual decision. A general mechanism has been suggested to involve the temporal integration and comparison of the outputs of different pools of sensory neurons in order to compute a decision variable (Heekeren, Marrett, Bandettini, & Ungerleider, 2004). This temporal integration might be performed in one of several regions such as the dorsolateral prefrontal cortex (DLPFC) (Heekeren et al., 2004; Kim & Shadlen, 1999), lateral intraparietal area (LIP) (Shadlen & Newsome, 2001), superior colliculus (SC) (Horwitz & Newsome, 1999), frontal eye fields (FEF) (Schall, 2002; Schall & Thompson, 1999; Thompson, Hanes, Bichot, & Schall, 1996) or intraparietal sulcus (IPS) (Colby & Goldberg, 1999), which all cooperate in order to translate the accumulated evidence into an action (Heekeren, Marrett, & Ungerleider, 2008; Rorie & Newsome, 2005). Neuronal activity in integrator areas gradually increases and then remains elevated until a response is given, with the rate of increase being slower during more difficult trials. A successful approach to explaining these kinds of neurophysiological data has been through the means of drift-diffusion or race models (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Schall & Thompson, 1999; Smith & Ratcliff, 2004), in which the noisy sensory information is integrated over time until a decision threshold is reached.

Here we present a large-scale model of a hierarchical spiking neural network (SNN) that integrates a low-level memory encoding mechanism with a higher-level decision process to perform a visual classification task in real-time. The model consists of Izhikevich neurons and conductance-based synapses for realistic approximation of neuronal dynamics (Dayan & Abbott, 2001; Izhikevich, 2003; Izhikevich, Gally, & Edelman, 2004), a STDP synaptic learning rule with additional synaptic dynamics for memory encoding (Brader et al., 2007), and an accumulator model for memory retrieval and categorization (Smith & Ratcliff, 2004). Grayscale input images were fed through a feed-forward network consisting of V1 and V2, which then projected to a layer of downstream classifier neurons through plastic synapses that implement the STDP-like learning rule mentioned above. Population responses of these classifier neurons were then integrated over time to make a perceptual decision about the presented stimulus. The full network, which comprised 71,026 neurons and approximately 133 million synapses, ran in real-time on a single off-the-shelf graphics processing unit (GPU).

In order to evaluate the feasibility of our model, we applied it to the extensively studied MNIST database of handwritten digits (LeCun et al., 1998). Due to the large variability within a given class of digits and a high level of correlation between members of different classes, the database provides stimuli whose categorization might span a wide range of difficulty levels, and as such is well-suited as a first benchmark for our model. However, it should be noted that MNIST does not pose many of the challenges of biological vision, such as distractors, occluders or translation invariance. Moreover, all the images are static and isolated in their receptive field. The network achieved 92% correct classifications, which is comparable to other SNN approaches (Brader et al., 2007; Querlioz, Bichler, & Gamrat, 2011) and simple machine learning algorithms (such as linear classifiers, k-Nearest Neighbor classifiers and simple artificial neural networks LeCun et al., 1998), but not to state-of-the-art models whose performance is close to 99.8% (Ciresan et al., 2011; Niu & Suen, 2012).

Additionally, our network produces reaction time (RT) distributions that are comparable to the behavioral RT distributions reported in psychophysical experiments. For example, we show that when the network makes an error, its RT is significantly slower than when making a correct class prediction; and that RTs do not decrease when the target stimulus has become familiar, which has also been observed in a rapid categorization study (Fabre-Thorpe, Richard, & Thorpe, 1998).

Although the present model does not reach the performance of specialized classification systems (Ciresan et al., 2011; Niu & Suen, 2012), our model represents a first step towards the construction of a general-purpose neurobiologically inspired model of visual recognition and perceptual decision-making. The model includes many neurobiologically inspired details not found in the algorithms described above. The present network was constructed on a publicly available SNN simulator that uses design principles, data structures, and process flows that are in compliance with general-purpose neuromorphic computer chips, and that allows for real-time execution on off-the-shelf GPUs (Richert, Nageswaran, Dutt, & Krichmar, 2011); its neuron model, synapse model, and address-event representation (AER) are compatible with recent neuromorphic hardware (Srinivasa & Cruz-Albrecht, 2012). Because of the scalability of our approach, the current model can readily be extended to an efficient neuromorphic implementation that supports the simulation of more generalized object recognition and decision-making regions found in the brain. Ultimately, understanding the neural mechanisms that mediate perceptual decision-making based on sensory evidence will further our understanding of how the brain is able to make more complex decisions we encounter in everyday life (Lieberman, 2007), and could shed light on phenomena like the misperception of objects in neuropsychiatric disorders such as schizophrenia (Persaud & Cutting, 1991; Summerfield, Egner, Mangels, & Hirsch, 2006).

2. Methods

We performed all simulations in a large-scale SNN simulator which allows the execution on both generic x86 central processing units (CPUs) and standard off-the-shelf GPUs (Richert et al., 2011). The simulator provides a PyNN-like environment (PyNN is a common programming interface developed by the neuronal simulation community) in C/C++ and is publicly available at <http://www.socsci.uci.edu/~jkrichtma/Richert-FrontNeuroinf-SourceCode.zip>. The simulator's API allows for details and parameters to be specified at the synapse, neuron, and network levels. In our experiments we used an Intel Xeon X5675 CPU operating at 3.07 GHz and a single NVIDIA Tesla M2090 GPU card.

All parameter values can be found in Table 1.

Download English Version:

<https://daneshyari.com/en/article/406408>

Download Persian Version:

<https://daneshyari.com/article/406408>

[Daneshyari.com](https://daneshyari.com)