Familiarity Detection in Cortical Microcircuits: Analysis

Overview

The research paper by Zhang et al. (2017), published in eNeuro, investigates how cortical microcircuits with NMDA receptor (NMDAR)-dependent bidirectional synaptic plasticity can detect familiar sensory inputs (e.g., face images) through unsupervised learning. The study uses Liquid State Machine (LSM) simulations in CSIM to model recurrent spiking neural networks, demonstrating familiarity detection as an intrinsic property of these circuits. Here I summarizes the paper's neuron types, connectivity patterns, and computational functions, and explains my understanding of the research.

Neuron Types

The LSM network includes:

- **Excitatory Neurons** (80%): Leaky Integrate-and-Fire (LIF) neurons with NMDAR synapses, supporting long-term potentiation (LTP), long-term depression (LTD), and spike-timing-dependent plasticity (STDP). Dynamics follow: t * dVm/dt = -(Vm - Vresting) + Rm * (Isyn + Iinject + Inoise) Parameters: t = 30 ms, Vresting = 0 mV, Rm = 1 MOhm, threshold Vth = 15 mV, reset potential in [-1, 1] mV, refractory period 3 ms. Synaptic weights are plastic, bounded by [1.0 * 10^-9, 6.5 * 10^-8].
- **Inhibitory Neurons** (20%): LIF neurons with static synapses, negative weights, and a 2 ms refractory period for network stability.
- **Input Neurons**: Excitatory neurons delivering spike trains based on image pixel intensities, with fixed weights (2.7 * 10^-7 for small networks) or gamma-distributed weights.

Connectivity Patterns

The network connectivity mimics cortical microcircuits:

- **Network Structure**: Neurons are arranged on a 3D grid (e.g., 10 * 10 * 5, 50 * 50 * 6) with recurrent connections.
- **Connection Probability**: Distance-dependent, defined as: P(D) = C * exp(-D^2(a, b) / l^2) where D(a, b) is the Euclidean distance, C is connection strength (base: 0.3 for EE, 0.2 for EI, 0.4 for IE, 0.1 for II, modulated by Cscale), and l controls connection length (e.g., l = 2.0 or l = infinity for random connections).
- **Synaptic Weights**: Initial weights follow a gamma distribution: f(x | a, b) = (1 / (b^a * G(a))) * x^(a-1) * e^(-x/b), a = 1 / SH_W^2, b = W * SH_W^2 with base W values (3 * 10^-8 for EE, 6 * 10^-8 for EI, -1.9 * 10^-8 for IE/II) modulated by Wscale (e.g., 0.9). Excitatory synapses are plastic; inhibitory and input synapses are static.

Computational Functions

The LSM performs:

- **Familiarity Detection**: Increases firing rates for familiar images post-learning, forming stimulus-specific subnetworks via NMDAR plasticity.

- **Plasticity**: Calcium-dependent plasticity follows: $dCa(t)/dt = I_{NMDA}(t) - (1 / t_{Ca}) * [Ca(t)]$ where I_{NMDA} depends on EPSPs and backpropagating action potentials. High $[Ca] > a2$ induces LTP; $a1 < [Ca] < a2$ induces LTD, enabling one-shot learning.
- **Generalization**: Generalizes familiarity to novel inputs of the same class (e.g., novel faces).
- **Storage Capacity**: Large networks ($50 * 50 * 6$, 15,000 neurons) store up to 60 familiarity traces via overlapping subnetworks.
- **Analysis**: Fisher's Discriminant Ratio: $J(t) = (\sum_C (\mu_C(t) - \mu(t))^2) / (\sum_C \sum_{i \in C} (S_i(t) - \mu_C(t))^2)$ identifies critical neurons, with deep-layer neurons recruited for familiarity encoding.

What I Understood from the Research Paper

The research paper by Zhang et al. (2017) offers a profound insight into the brain's capacity to recognize familiar sensory inputs, such as faces, through an inherent mechanism within cortical circuits. My understanding is that the study demonstrates how these circuits, modeled as a network of interconnected neurons, can identify and recall sensory patterns without requiring explicit training. This capability is central to how our brains process and retain information, enabling us to recognize familiar objects or people effortlessly.

The paper's model, a Liquid State Machine, simulates a cortical microcircuit with two main types of neurons: excitatory (80%) and inhibitory (20%). Excitatory neurons use a Leaky Integrate-and-Fire mechanism, which governs their electrical activity, and possess adaptable connections driven by NMDA receptors. These receptors allow connections to strengthen (long-term potentiation) or weaken (long-term depression) based on calcium levels in the neuron. This adaptability enables the network to learn a stimulus, like an image, after a single exposure, producing a stronger response when the same stimulus reappears. Inhibitory neurons, with fixed connections, maintain network stability by balancing excitation.

A significant finding is the network's ability to generalize. After learning one face, it can recognize other faces of the same category (e.g., human faces) as somewhat familiar, reflecting a cognitive ability to extend knowledge to new instances. The model organizes neurons into a 3D grid, with connections between them following a distance-dependent rule: $P(D) = C * \exp(-D^2(a, b) / l^2)$. This structure allows larger networks, containing up to 15,000 neurons, to store as many as 60 memory traces by forming overlapping groups of neurons, each tuned to a specific familiar input. This organization ensures the network can learn new patterns while preserving existing memories.

The study converts images into patterns of neural activity through an input layer, which feeds into the main circuit. The network's response—higher activity for familiar inputs—is analyzed using a method called Fisher's Discriminant Ratio, which identifies neurons critical for distinguishing familiar from unfamiliar inputs. Notably, neurons in deeper layers of the network become more involved over time, suggesting a hierarchical processing approach akin to cortical function.

Comparing this to my own simulation, I recognize that the paper's model is tailored for a specific task—detecting familiar patterns—while my 500-neuron network simulates general brain activity . The paper's use of a larger network and sophisticated adaptability mechanisms highlights its ability to handle complex cognitive tasks, unlike my model's broader, less specialized dynamics.

In essence, the paper reveals that the brain's ability to recognize familiar inputs is a natural outcome of its circuitry, driven by adaptable connections in excitatory neurons. This insight underscores the efficiency and flexibility of cortical microcircuits in processing sensory information and forming memories, providing a valuable model for understanding biological cognition.

Link to the paper