CCN Methods

* The network was trained to classify 3 different letters that were presented in 4 quadrants
* To ensure competition between the quadrants, a convolutional kernel of the size of one quadrant, with 68 output nodes (68x28x28) was applied to the 56x56 input with a stride of 28 (moving over quadrants), resulting in an output of size 68x2x2 the activations in each quadrant were then summed up using a 2x2 sliding kernel, resulting in a hidden layer of size 68x1, connected to an output layer with 3 units
* Activationconvolutional kernel
* pre-activation
* shifted sigmoid with a set bias
* The shifted, steep sigmoid ensured sparse activations between 0 and 1, limiting the activations in the linear part of the sigmoid
  + This was important for the dynamics in the hidden layer described below

Dynamics

* τH : slope dynamics H
* τH : slope dynamics Θ
* H(h): activation in hidden node h
* Θ(h): relaxation corresponding to hidden node h
* a: scaling factor for relaxation (scaling = a\*max(Z))
* α(t): pulsed inhibition
* T < 1 pushes activations to 1
* To mimic gain modulation by attention, one input was multiplied with 1.1 (attended input) and the unattended input was multiplied with 0.9 (unattended input)

Write the same thing with optuna

* The dynamics required the activations in the hidden layer to be sparse, and with only few activations in the linear part of the sigmoid
* Sparsity constraint, whereby the bias term is penalized based on the difference between the activations and a set value rho ~= 0
* Sparsity formula:
* Moreover, we aimed to minimize the difference between the sum of the hidden activations of two individual inputs and the hidden activations to the sum of the two inputs
* Using optuna we identified the parameters for beta, rho, and eta that minimized the average loss over the last 20 out of 80 training epochs as well as ω

CCN Results

Figure

1. Network architecture
2. Network dynamics to one input
3. Network dynamics simultaneous inputs

* The network organises the competing inputs in accordance with their input gain and reads them out in succession, within the down-state of the pulsed alpha inhibition

Discussion

* The network implements multiplexing in neural network
* This is just proof of principle, but we are hoping to constrain these networks to electrophysiological data