Theoretical Background

**Dual-Process Signal Detection theory**

Familiarity is a type of recognition that gives a quantitative measure about a previously learned stimulus. That is, familiarity corresponds to the degree of similarity between the characteristics of a perceived stimulus and the characteristics of an old stimulus stored in memory. When familiar, a specific stimulus has a higher level of familiarity than a novel stimulus.

**Perirhinal cortex & novelty neurons**

The perirhinal cortex (PRC, in blue) seems to be crucial in the familiarity processes:

- Novelty neurons in the PRC respond stronger when presented with a new stimulus. Once familiar, the activity of novelty neurons in the PRC is reduced.

Modeling with Hebbian learning

Artificial neural networks successfully used Hebbian learning to model familiarity in the PRC on formal binary patterns:

- Familiar stimuli have more inhibition than novel ones.

Modeling & Methodology

Simulations took place according to Standing’s experiment. Three main simulations were performed to explore:

- The memory capacity of the model
- The presence of recency/primacy effects
- The performances on similar images (e.g., only cat images)

Simulation 1 showed that the Memory Capacity of the model is up to 40 images. Error probability and the number of images retained (N_e) were computed over the entire task. Experimental data are depicted in black. More efficient CNN improves model performances (AlexNet > ResNet50).

Simulation 2 showed a Recency-like Effect when the number of learned images did not exceed the memory capacity of the model. No Primacy Effect was observed.

Simulation 3 showed a Similarity Effect. Model performances collapse when trained with images only from one semantic category (N = 40).

Performance is worse when tested with cats than with dogs, thus showing that the model is sensitive to homogeneity between inputs during training.

Discussion

Our results suggest the Hebbian model learns the global representation of a stimulus. This could be done by encoding correlations shared by several stimuli. Here, familiarity comes when the representation matches with the stimulus during the testing phase. Interestingly, model performances decrease drastically when too many stimuli are presented as well as when there is high homogeneity between inputs. It is then plausible that the model learns patterns with not enough details to allow discrimination. This would predict the high overlapping observed between d curves when increasing the dataset size.

Our computations are consistent with global matching theories of familiarity. However, it is known that familiarity could arise from different coexisting mechanisms (e.g., stimulus-specific reduction of neural activity by repetition suppression). Nevertheless, it is not clear what are the conditions for a specific mechanism to take precedence over another. Here, we posit that familiarity could be expressed through the overall structure of stimuli (i.e., global matching) if the amount of information to be encoded is limited. Beyond a certain threshold of information to be learned, we hypothesize that familiarity would be expressed through other mechanisms.

References