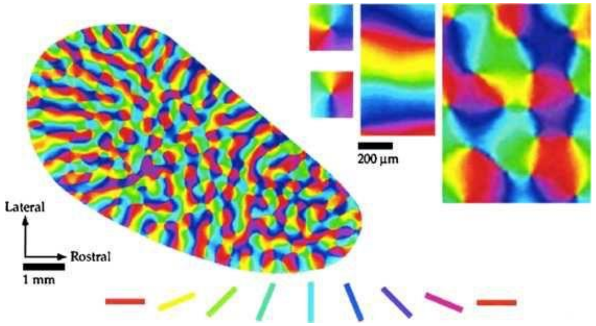
Credit-based self-organizing maps: training deep topographic networks with minimal performance degradation

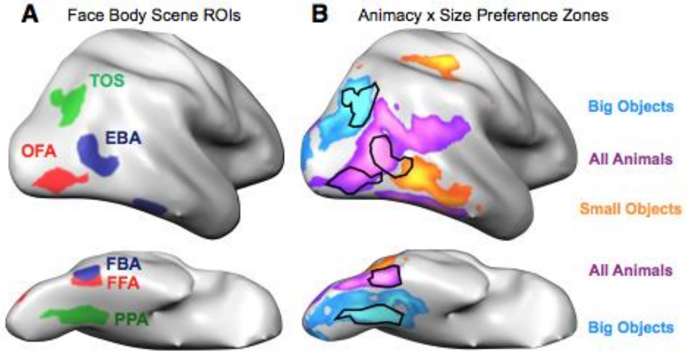
Amir Ozhan Dehghani, Xinyu Qian, Asa Farahani, Pouya Bashivan

Motivation

- **The brain is topographically organized**: In humans and other animals, functionally similar neurons are spatially close to each other
- The current driving mechanism of topographical organization is unknown.
- We sought to understand the computational principles underlying visual topographical organizations using Deep Neural Networks.



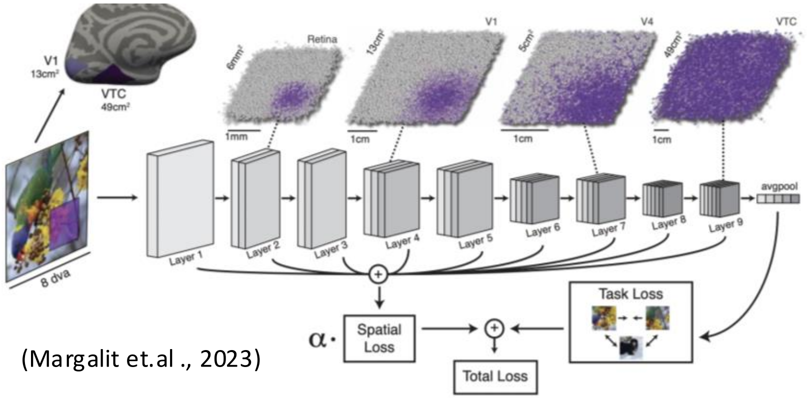
(Bosking., 1997)



(Konkle et al., 2013)

Prior Work

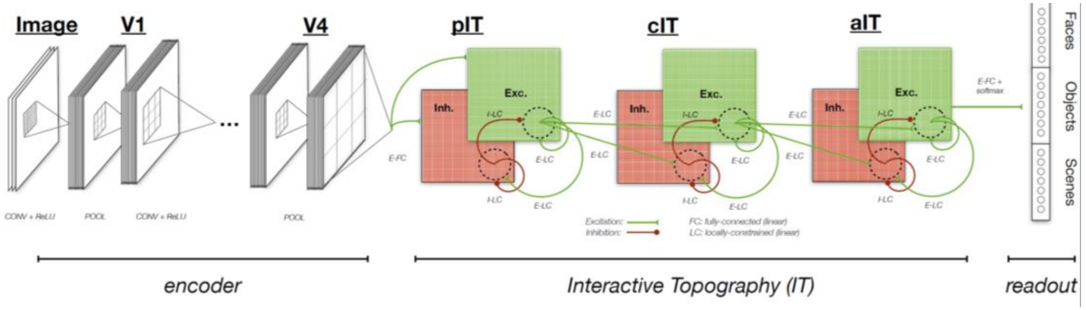
- Previous approaches to simulate topographical organization in DNNs involved using objective functions.



(Margalit et.al ., 2023)

Topographic Deep Artificial Network (TDANN)

- TDANN uses an objective function that forces units' responses to be correlated as a function of their distance



(Blauch et.al ., 2022)

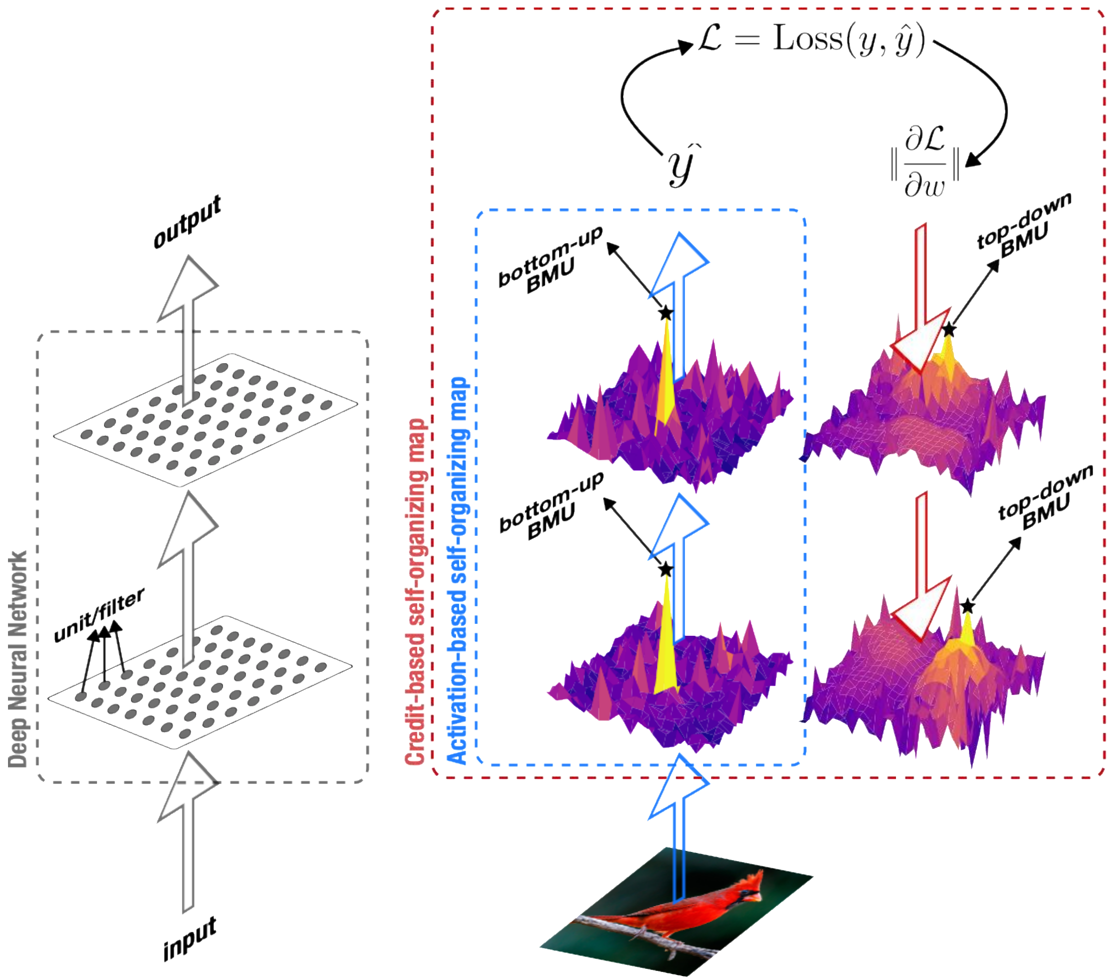
Interactive Topographic Network (ITN)

- ITN uses an objective function that penalizes long-distance connections

However, this approach often resulted in weaker performance on categorization

Our Proposal

- We propose an alternative approach based on self-organizing maps: ***Credit-Based Self-Organizing Maps (CB-SOM).***
- Unlike the classical Kohonen's SOM, the BMU is chosen based on unit activation. CB-SOM chooses BMU based on the gradient to the objective function.
- This results in the selection of the unit with the ***highest impact*** on the object categorization performance.



SOM Update Rule

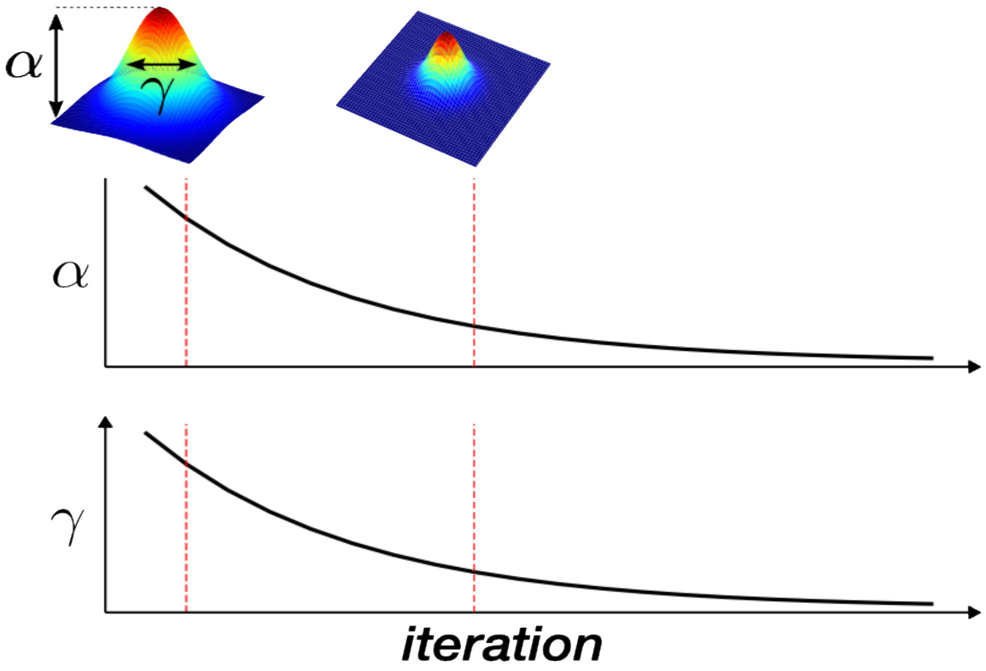
- We enforce topographical organization by:
  - We assign 2D positional coordinates for every unit in the convolutional layer to simulate a cortical sheet
  - A best-matching unit is chosen based on the credit-based or activation-based method.
  - A neighbourhood function is constructed based on the distance of units to the BMU coordinate

$$\gamma_{c,ij}(t) = \exp\left(-\frac{\|\mathbf{p}_c - \mathbf{p}_{ij}\|^2}{2\sigma(t)^2}\right)$$

  - Update weights for every unit layer using the SOM updating rule

$$\mathbf{w}_{ij}(t+1) = \mathbf{w}_{ij}(t) + \alpha(t) \cdot \gamma_{c,ij}(t) \cdot (\mathbf{w}_c(t) - \mathbf{w}_{ij}(t))$$

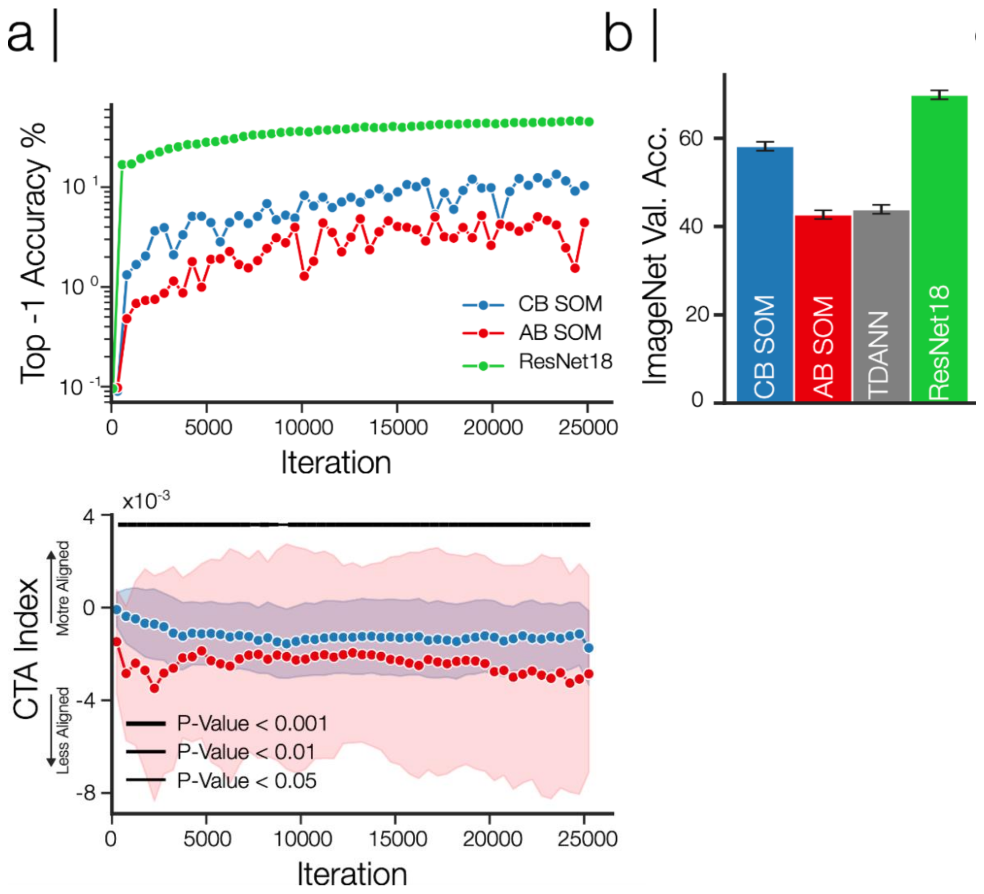
  - The two parameters α , γ decay exponentially



Results – Boost in Performance

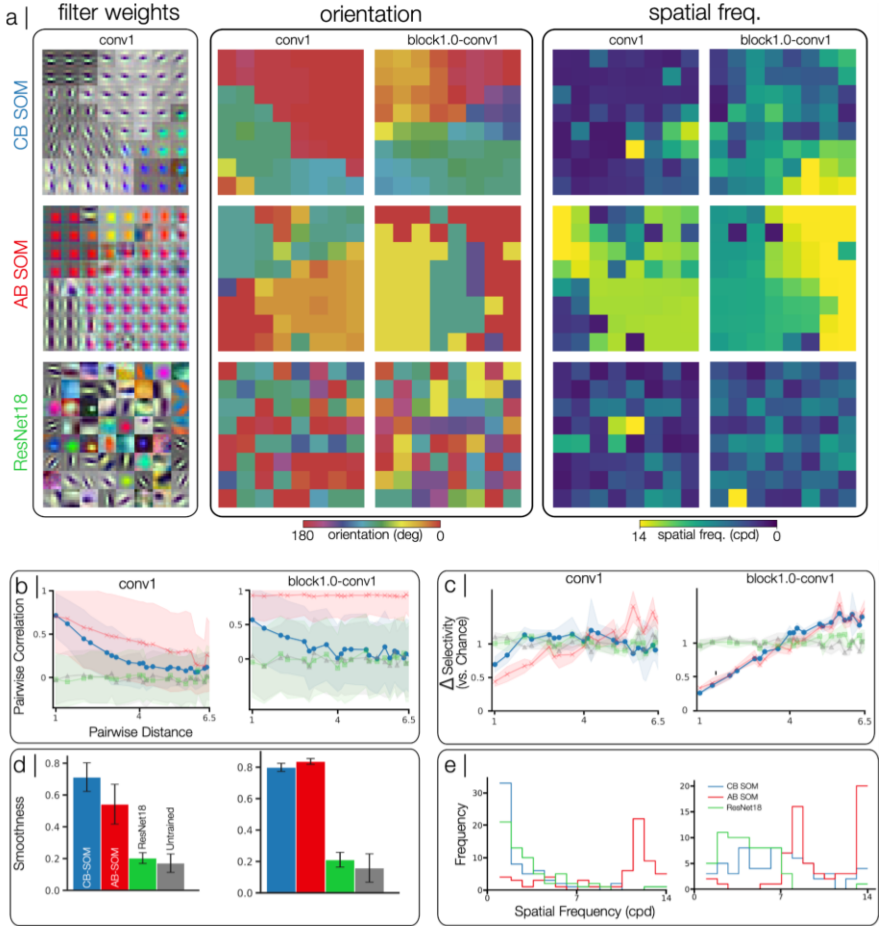
- Using our approach, the CB-SOM model significantly outperforms the AB-SOM model while preserving topographical organization.
- We quantified the impact of SOM updates on model performance in object categorization.

CB-SOM update is more aligned with the object recognition task



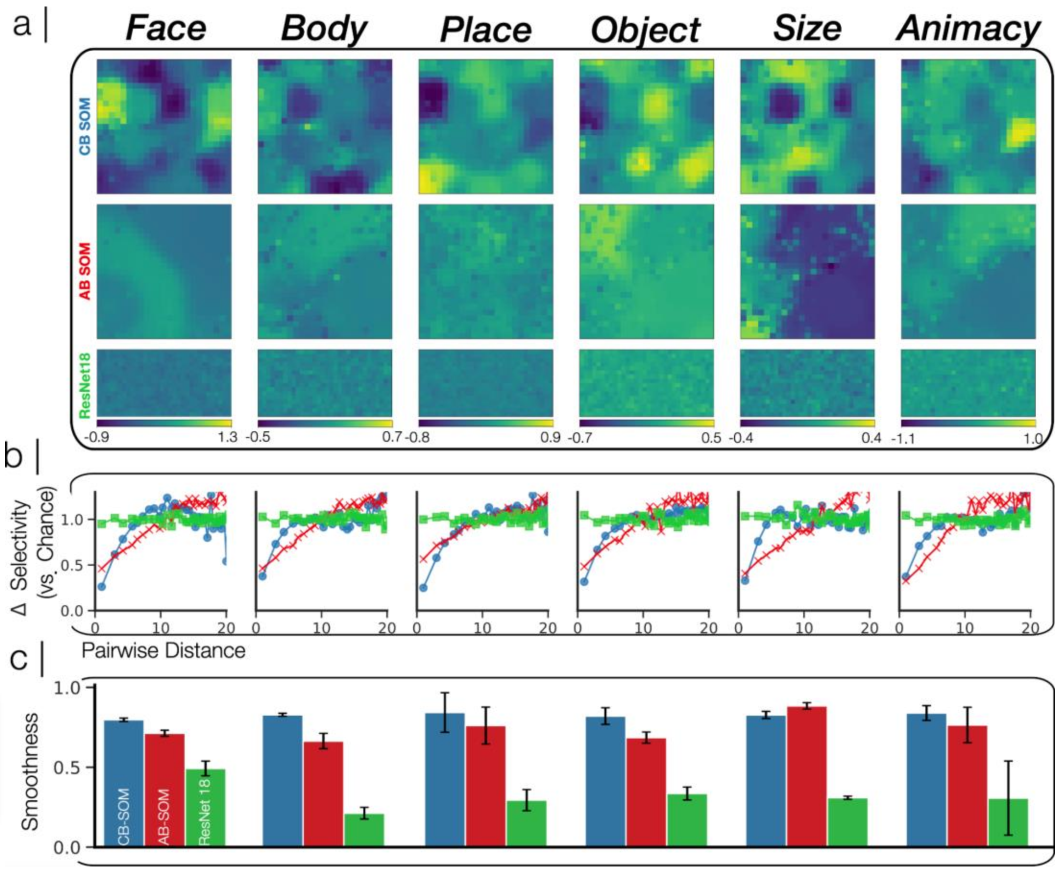
Results – Topographical Representation Primary Visual Cortex

- CB-SOM replicates the topographical organization in the primary visual cortex. (i.e. orientation and spatial frequency)
- We observed an exponential decay of pairwise correlation between filters as a function of Euclidean distance.
- Resulting in smooth transition between units selectivity



Results – Topographical Representation Inferior Temporal Visual Cortex

- CB-SOM replicates the topographical organization in higher-level visual regions according to category selectivity. (i.e. Face, Body regions)
- CB-SOM exhibits higher smoothness and category-selective regions.



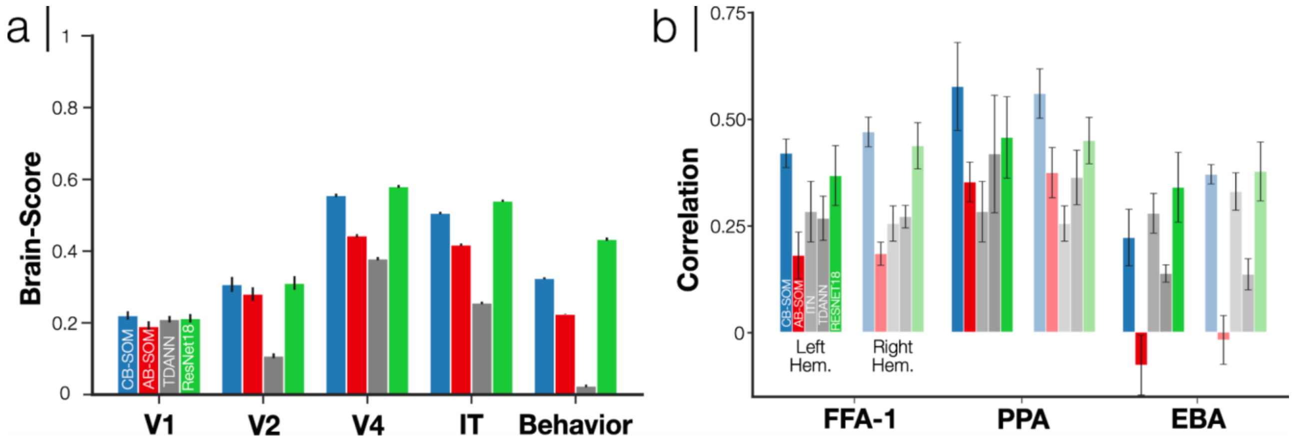
CB-SOM can replicate topographical organization in the early and late visual areas

Results – Model-to-Brain Similarity

- We compared the model's similarity to neural data for the visual cortex for macaques (Brain-Score) and humans (NSD)

- CB-SOM shows substantially higher alignment with neural activity in the visual cortex of both macaques (Brain-Score) and humans (NSD

CB-SOM can develop representations that closely resemble primates' neural representations



Discussion

- Top-down vs Bottom Up:
  - CB-SOM selects winner units based on their contribution to reducing a task-relevant objective function (top-down).
- Principles of topographical organization:
  - Our results indicate that topographical organization emerges not just from wiring cost minimization but also through self-organizing processes driven by task relevance.
- Universal topography across layers and models:
  - Unlike many models that focus on a single visual area or a limited range of layers, CB-SOM induces topography throughout all network layers with minimal loss in task performance for both ResNet18 and CorNet-S architectures.

Acknowledgment

Xinyu Qian

Asa Farahani

Dr. Pouya Bashivan

Amirozhan.dehghani@mail.mcgill.ca

@great_caster