Unsupervised neural network models of the ventral visual stream

Typical neural network models of the ventral stream are built via supervised training methods involving huge numbers of semantic labels. In particular, today’s best models of visual cortex are trained on ImageNet, a dataset that contains millions of category-labeled images organized into thousands of categories.

But, cannot provide a correct explanation of how such representations are learned in the first place.

* Here: Identifying unsupervised learning procedures that achieve good performance on challenging sensory tasks and effective predictions of neural responses in visual cortex would thus fill a major explanatory gap.

Contrastive methods:

* Instance recognition (IR)
* Contrastive Multiview coding (CMC)
* Momentum contrast (MoCo)
* Simple contrastive learning of representation (SimCLR)
* Local aggregation (LA)

To evaluate these unsupervised learning algorithms: trained using ResNet18 network architecture.

Semi-supervised learning algorithms:

* Local label propagation (LLP)
* Mean teacher (MT)