Mine your own view: A self-supervised approach for learning representations of neural activity

Mehdi Azabou[1], Mohammad Gheshlaghi Azar[2], Ran Liu[1], Chi-Heng Lin[1], Erik C. Johnson[4], Kiran Bhaskaran-Nair[3], Max Dabagia[1], Bernardo Avila-Pires[2], Lindsey Kitchell[4], Keith B. Hengen[3], William Gray-Roncal[4], Michal Valko[2], Eva L. Dyer[1,5]

1-Georgia Institute of Technology, 2-DeepMind, 3-Washington University in St. Louis, 4-Johns Hopkins University Applied Physics Laboratory, 5-Emory University

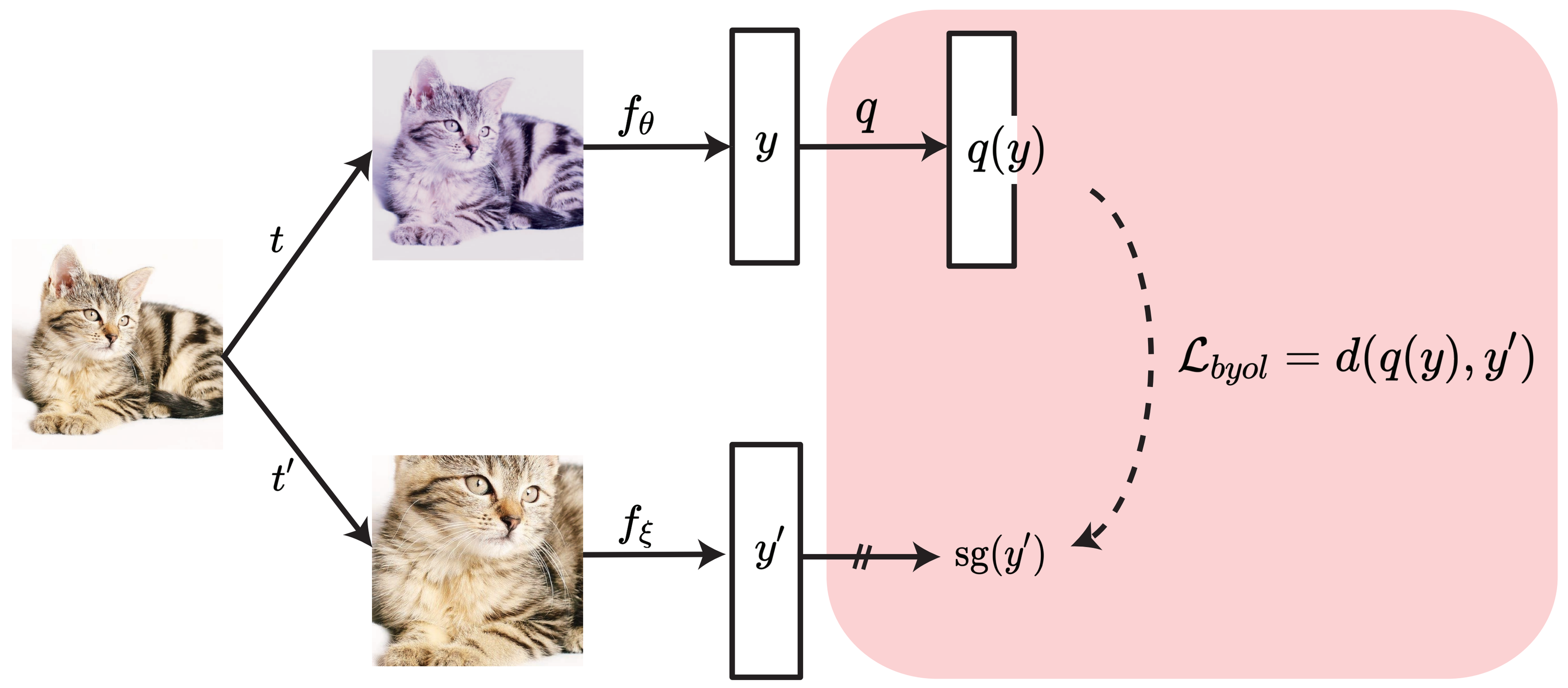
Motivation

Traditionally, neural decoding has been performed through supervised approaches that aim to map specific behaviors or stimuli to specific neural activity patterns through labeled data. However, the representations learned through a supervised approach typically require simple trial structure and repetitive behaviors, and fail to generalize to new datasets. Here, we ask whether we can use self-supervised learning principles to learn more robust and generalizable representations of neural activity.

Self-supervised Learning

Self-supervised methods leverage augmentations to **build invariances into the representation.**

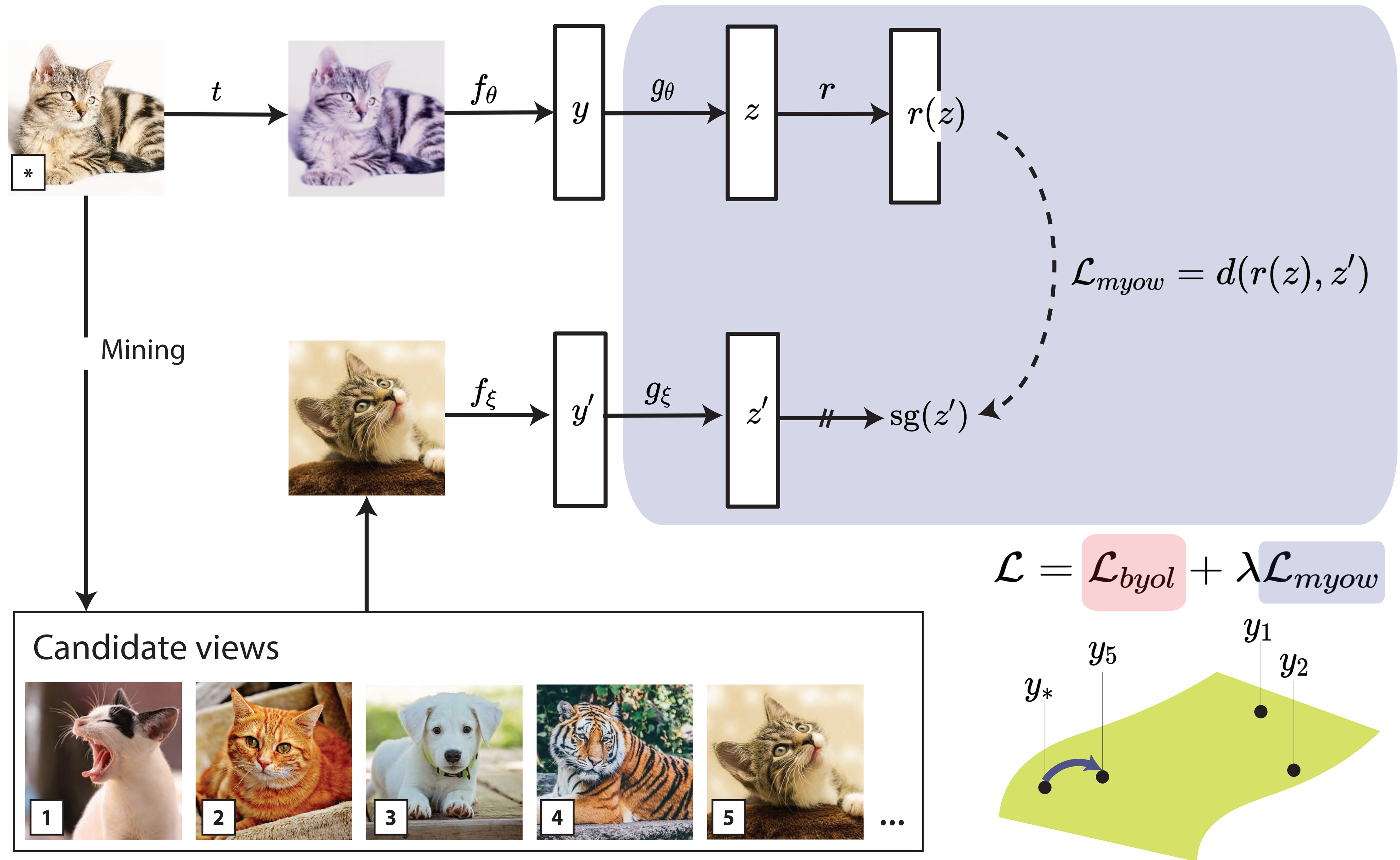
BYOL [1] does this by using a predictor tasked to predict across representations of two independently augmented views of the same sample.



Challenge:

Current self-supervised learning (SSL) methods depend on hand-crafted augmentations. What are the right augmentations for building robust and generalizable neural representations?

Approach



Mine the dataset for positive examples!

1. Feed the anchor view and candidate views through the online and target nets
2. Perform a kNN search in the representation space to find a positive view
3. After the cascaded projection, predict the projection of the mined view from that of the anchor view

MYOW for images

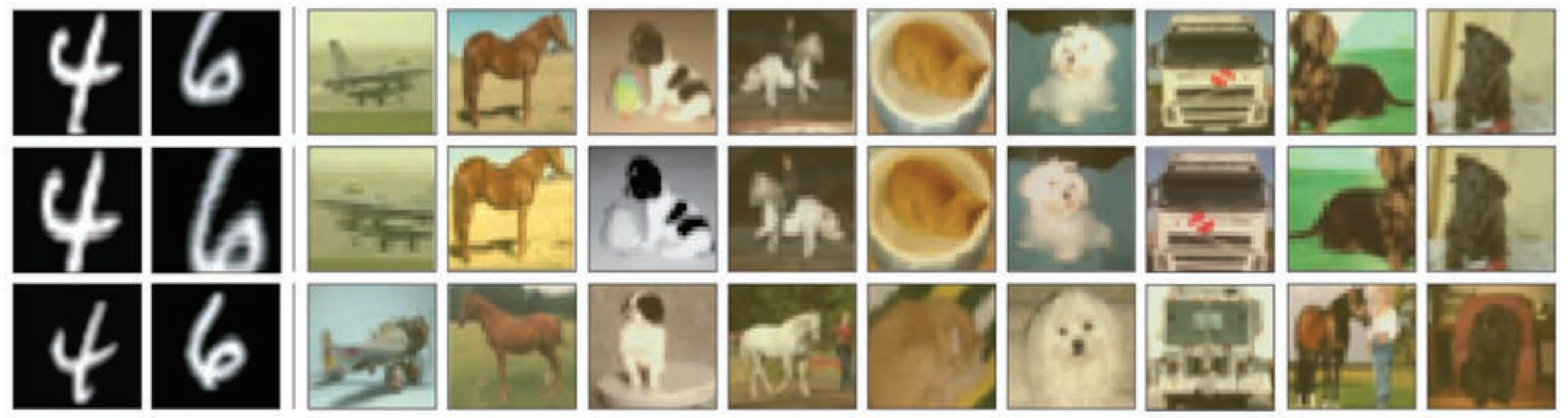
Method	ResNet-18 CIFAR-10	ResNet-18 CIFAR-100	ResNet-18 CIFAR-20	ResNet-18 Tiny ImageNet	ResNet-50 CIFAR-10	ResNet-50 CIFAR-100	ResNet-50 CIFAR-20
SimCLR *	91.80	66.83	-	48.84	91.73	-	-
BYOL	91.71	66.70	76.90	51.56	92.12	67.87	77.38
MYOW	**92.10**	**67.91**	**78.10**	**52.58**	**93.18**	**68.69**	**78.87**

Table 1 MYOW applied to different image datasets

Sample

Augmented view

Mined view



MYOW for neural decoding

We benchmark our approach, against standard methods, on datasets from trial-based reaching datasets from non-human primates, and on free behavior in rodent visual cortex and hippocampus.

	Chewie-1	Chewie-2	Mihi-1	Mihi-2	Rat-V1	Mouse-CA1
Supervised	63.29	72.29	63.64	61.49	86.34	93.01
pi-VAE	65.63	60.60	62.44	63.26	73.10	82.48
AE	48.40	46.79	50.94	55.19	34.17	57.73
RP	59.21	50.69	57.78	53.76	82.93	82.12
TS	60.16	49.48	59.23	54.10	82.45	81.93
SimCLR	61.36	51.62	59.41	56.29	81.03	81.94
BYOL	66.65	64.56	72.64	67.44	85.42	93.24
MYOW	**70.54**	**72.33**	**73.40**	**71.80**	**88.01**	**93.70**

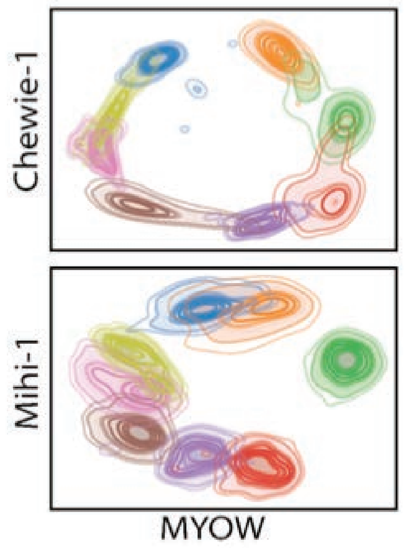


Table 2 Accuracy in the prediction of brain states from spiking neural activity. We considered the following decoding tasks (1) Reach direction decoding: Predict one of eight reach targets in the reaching task. (2) Arousal state decoding: Predict Rem, nRem, or Wake.

We show that by incorporating nonlocal but "similar" time points into the system, and predicting across these distinct time points, the network can build time-invariant representations that allow for more faithful decoding on downstream tasks and resilience to domain shift.

Website and code: https://nerdslab.github.io/myow/

References

[1] Grill et al., Bootstrap your own latent: A new approach to self-supervised learning, 2020

Funding

NIH BRAIN R01: Theory, Models, Methods Program (NIH- 1R01EB029852)