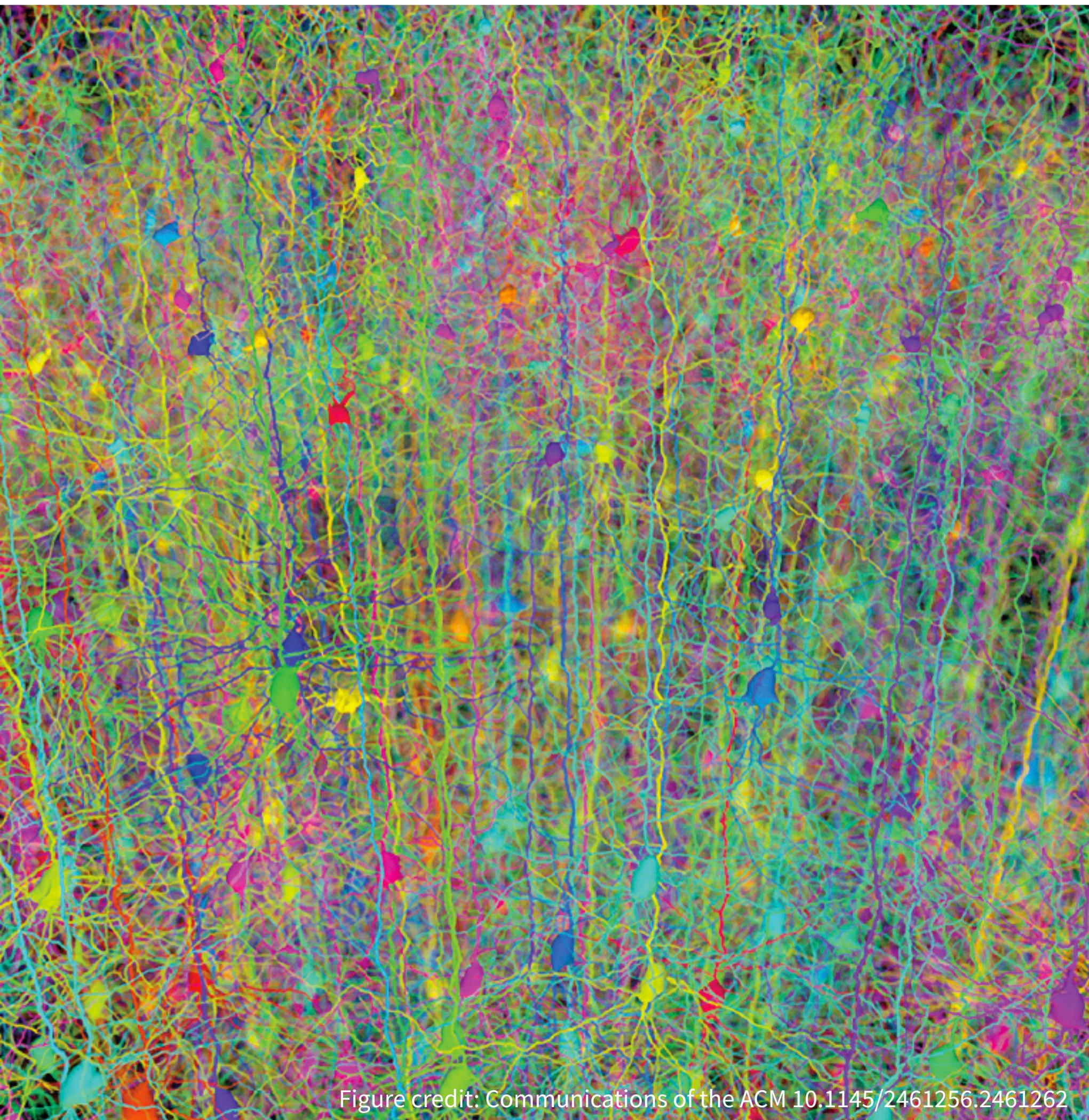


Unsupervised Learning

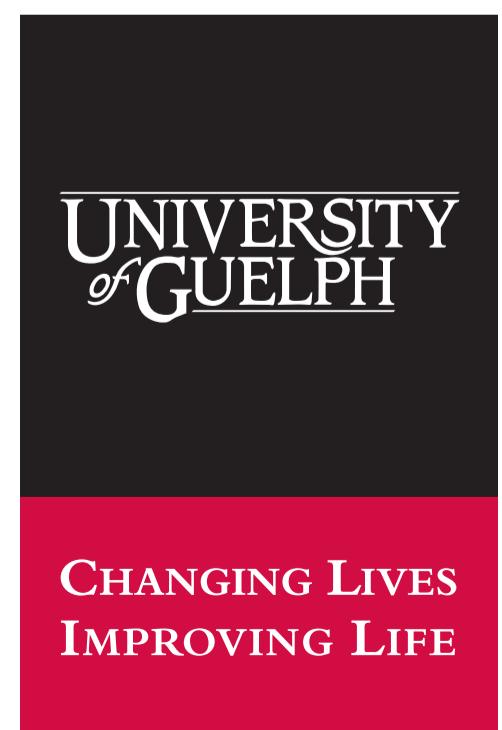


GRAHAM TAYLOR

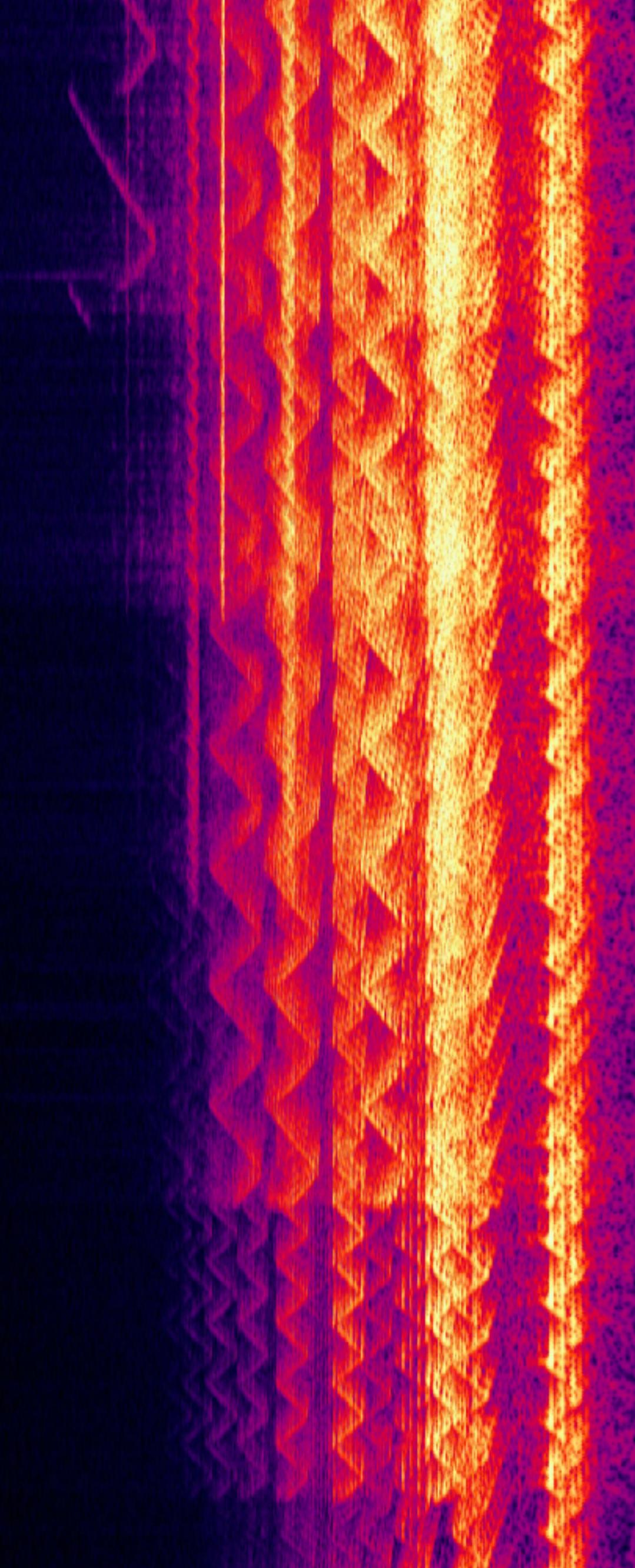
VECTOR INSTITUTE

SCHOOL OF ENGINEERING
UNIVERSITY OF GUELPH

CANADIAN INSTITUTE
FOR ADVANCED RESEARCH

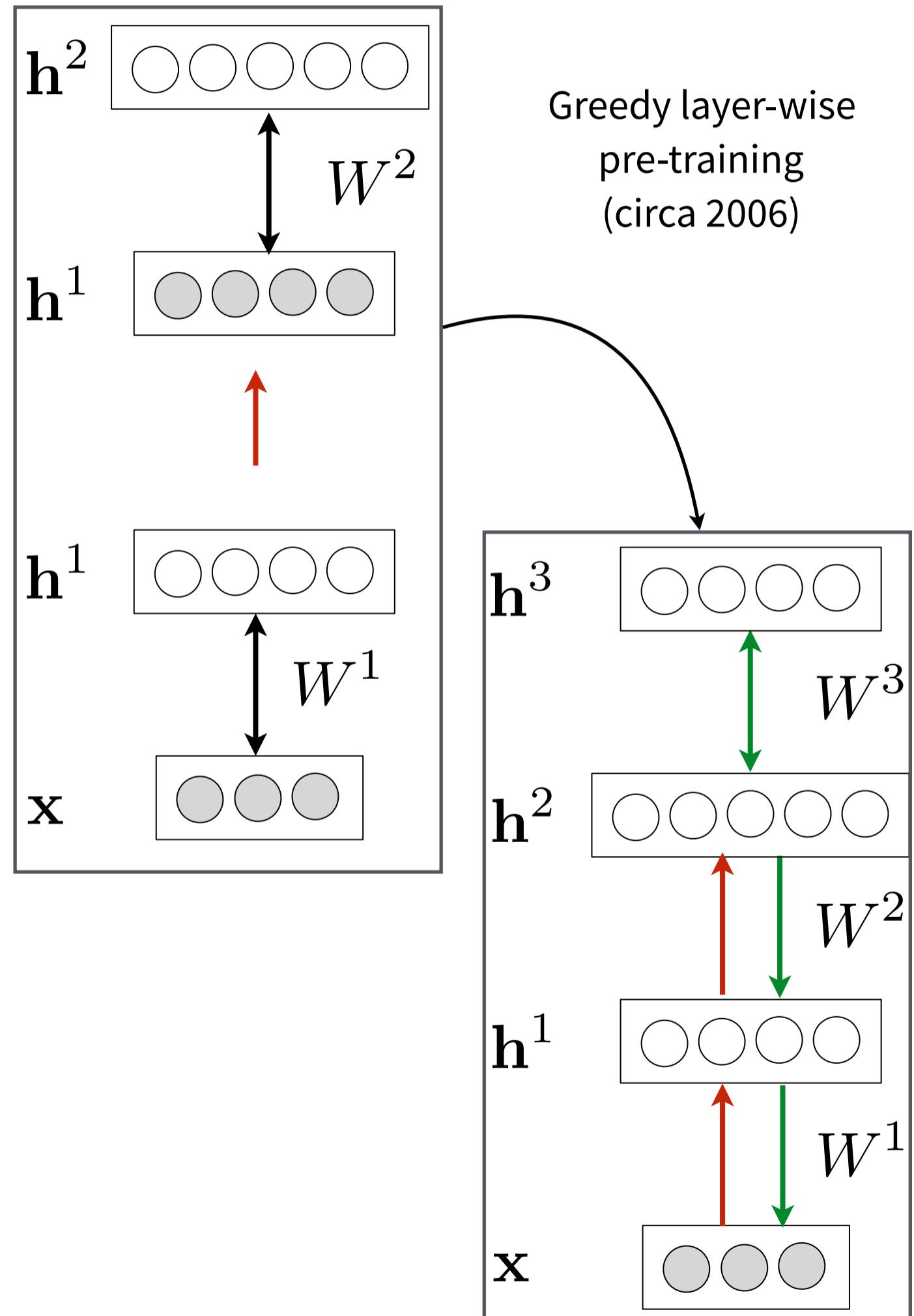


CIFAR
CANADIAN
INSTITUTE
FOR
ADVANCED
RESEARCH



An Interesting Historical Fact

- Unsupervised learning was the catalyst for the present DL revolution that started around 2006
- Now we can train deep supervised neural nets without “pre-training”, thanks to
 - Algorithms (nonlinearities, normalization, regularization)
 - More data
 - Better computers (e.g. GPUs)
- Should we still care about unsupervised learning?



Why Unsupervised Learning?

Reason 1:

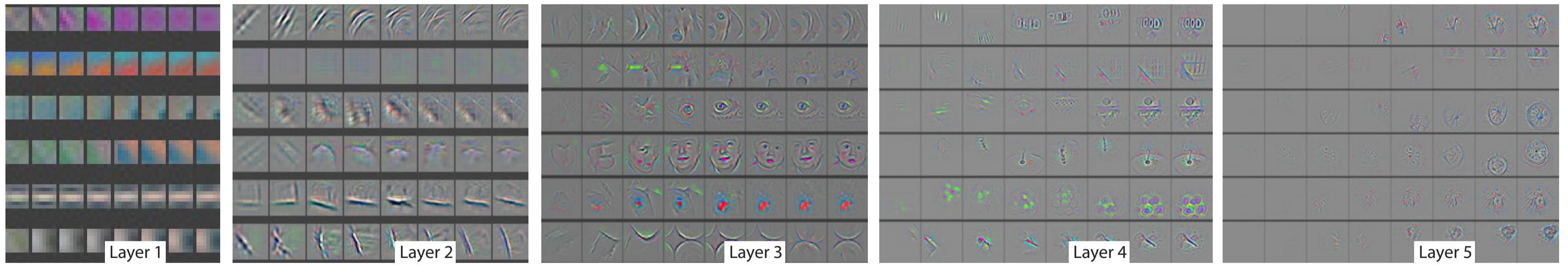
We can exploit unlabelled data; much more readily available and often free.



Why Unsupervised Learning?

Reason 2:

We can capture enough information about the observed variables so as to ask new questions about them; questions that were not anticipated at training time.



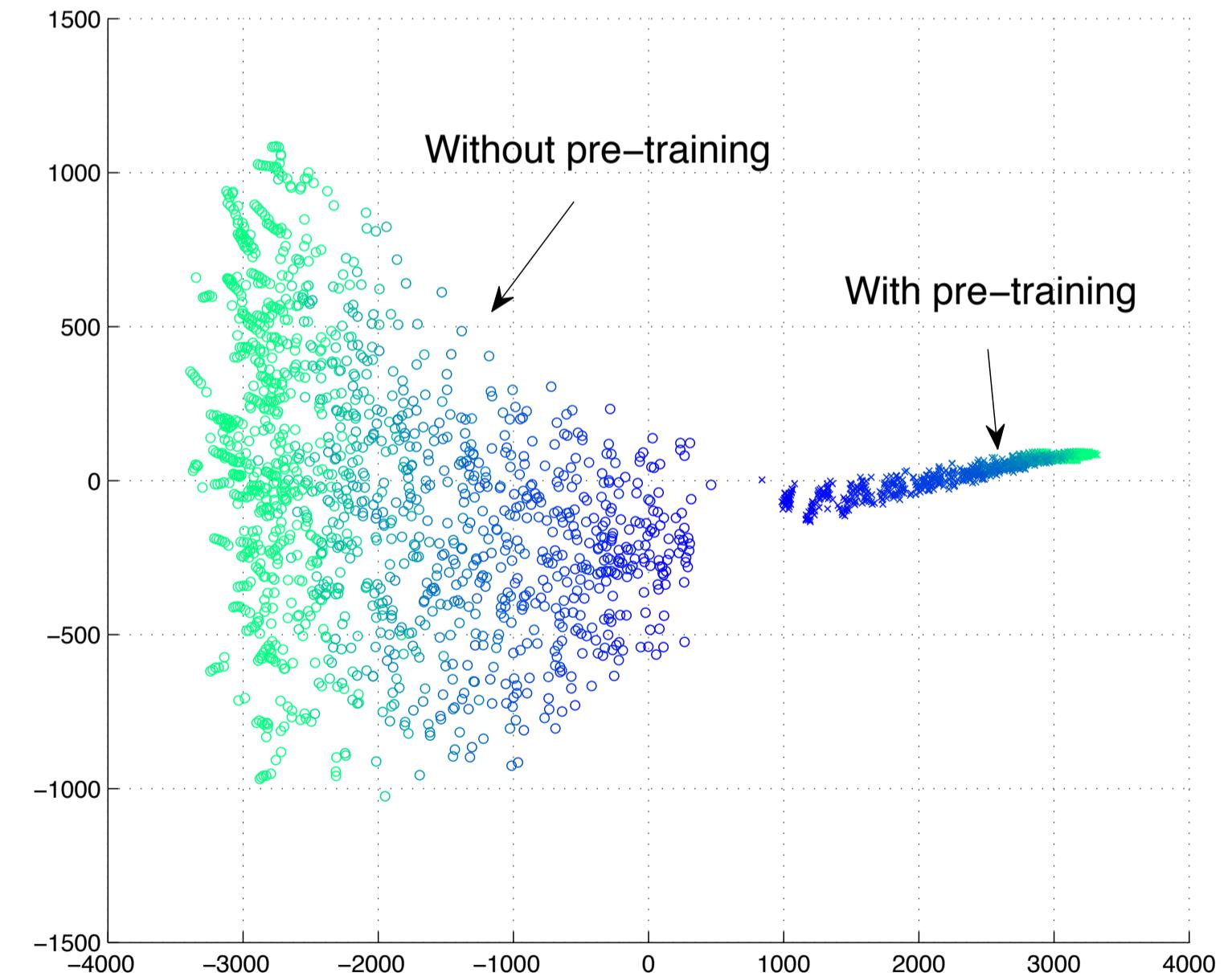
Why Unsupervised Learning?

Reason 3:

Unsupervised learning has been shown to be a good **regularizer** for supervised learning; it helps generalize.

This advantage shows up in practical applications:

- transfer learning,
domain adaptation
- unbalanced classes
- zero-shot, one-shot
learning



Why Unsupervised Learning?

Reason 4:

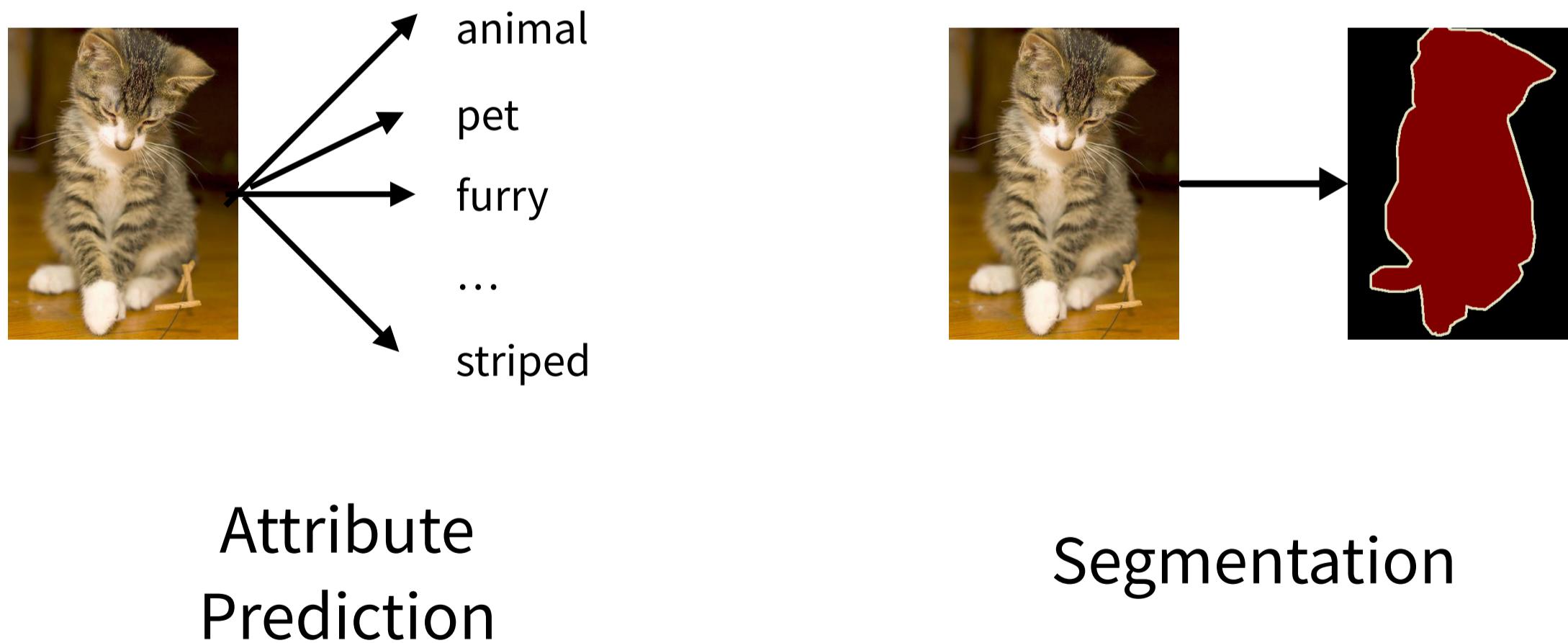
There is evidence that unsupervised learning can be achieved mainly through a level-local training signal; compare this to supervised learning where the only signal driving parameter updates is available at the output and gets backpropagated.



Why Unsupervised Learning?

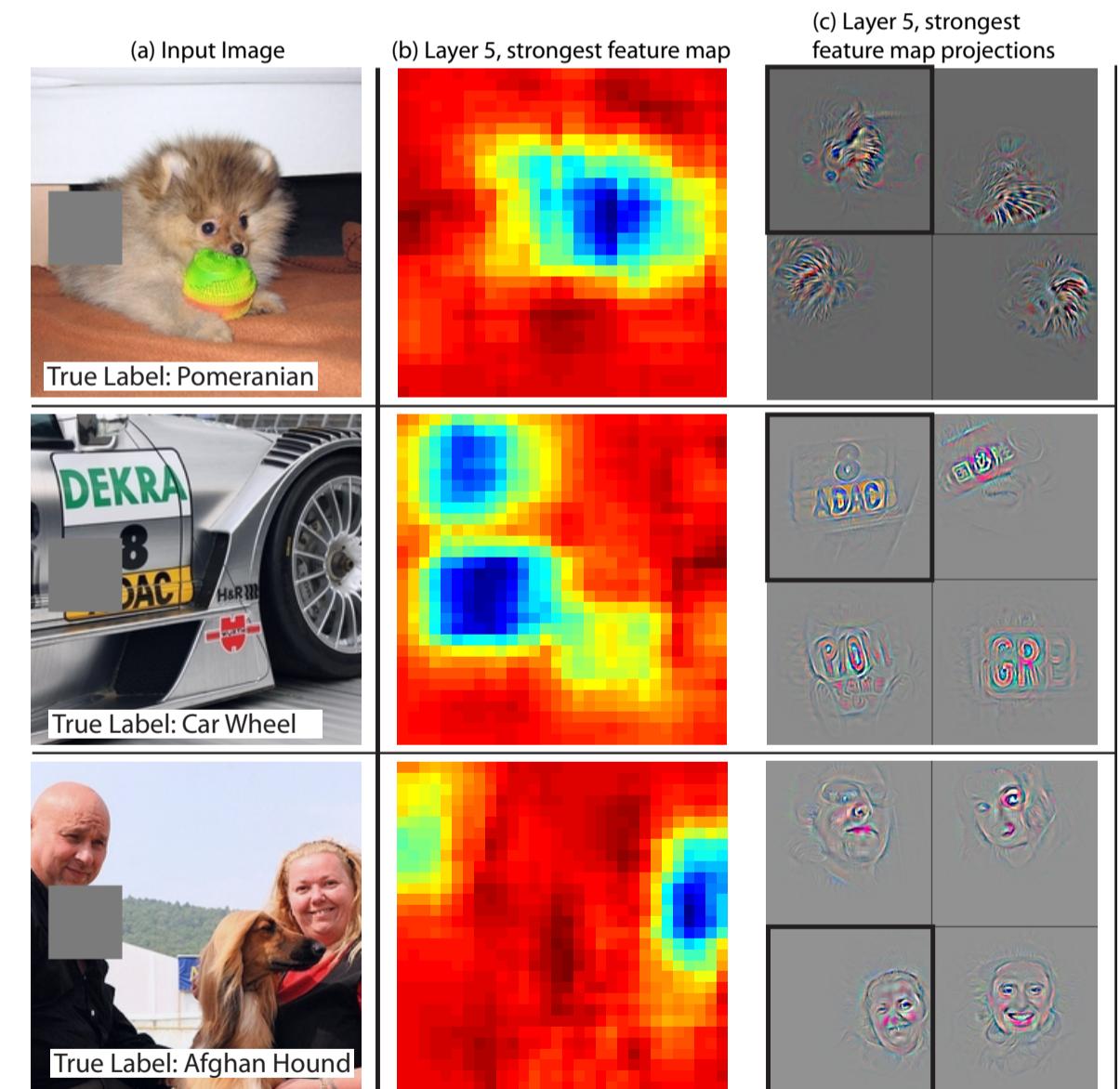
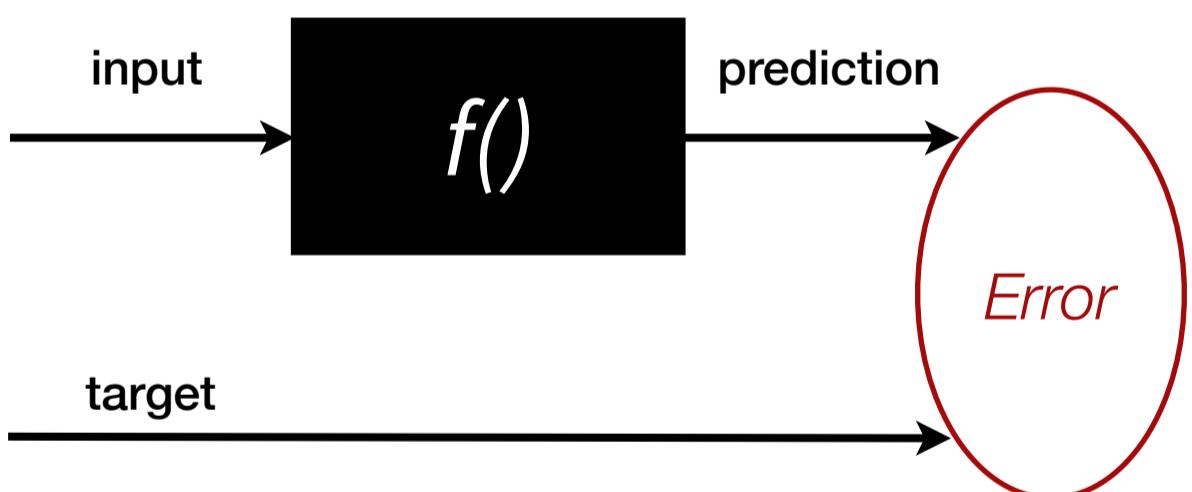
Reason 5:

A recent trend in machine learning is to consider problems where the output is high-dimensional and has a complex, possibly multi-modal joint distribution. Unsupervised learning can be used in these “**structured output**” problems.

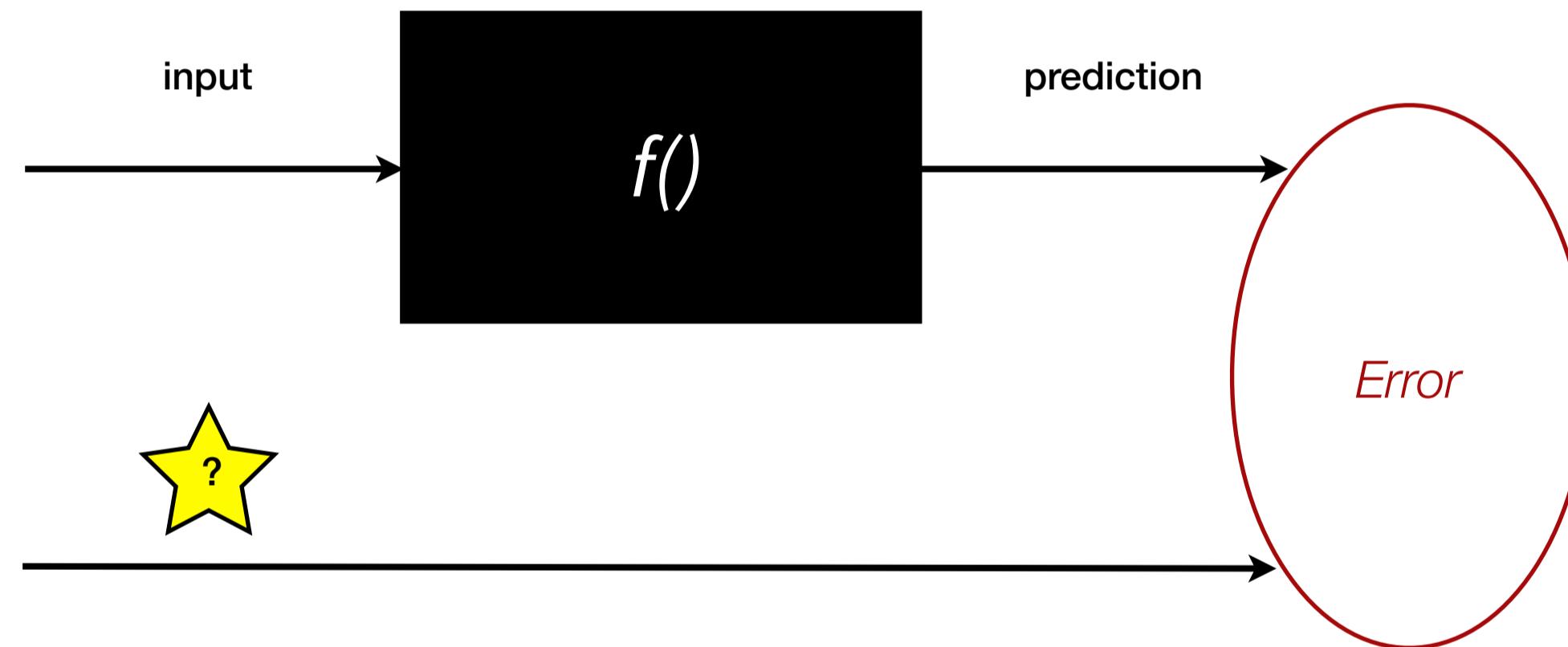


Supervised Learning of Representations

- Learn a representation with the objective of selecting one that is best suited for predicting targets given input

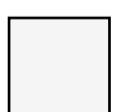
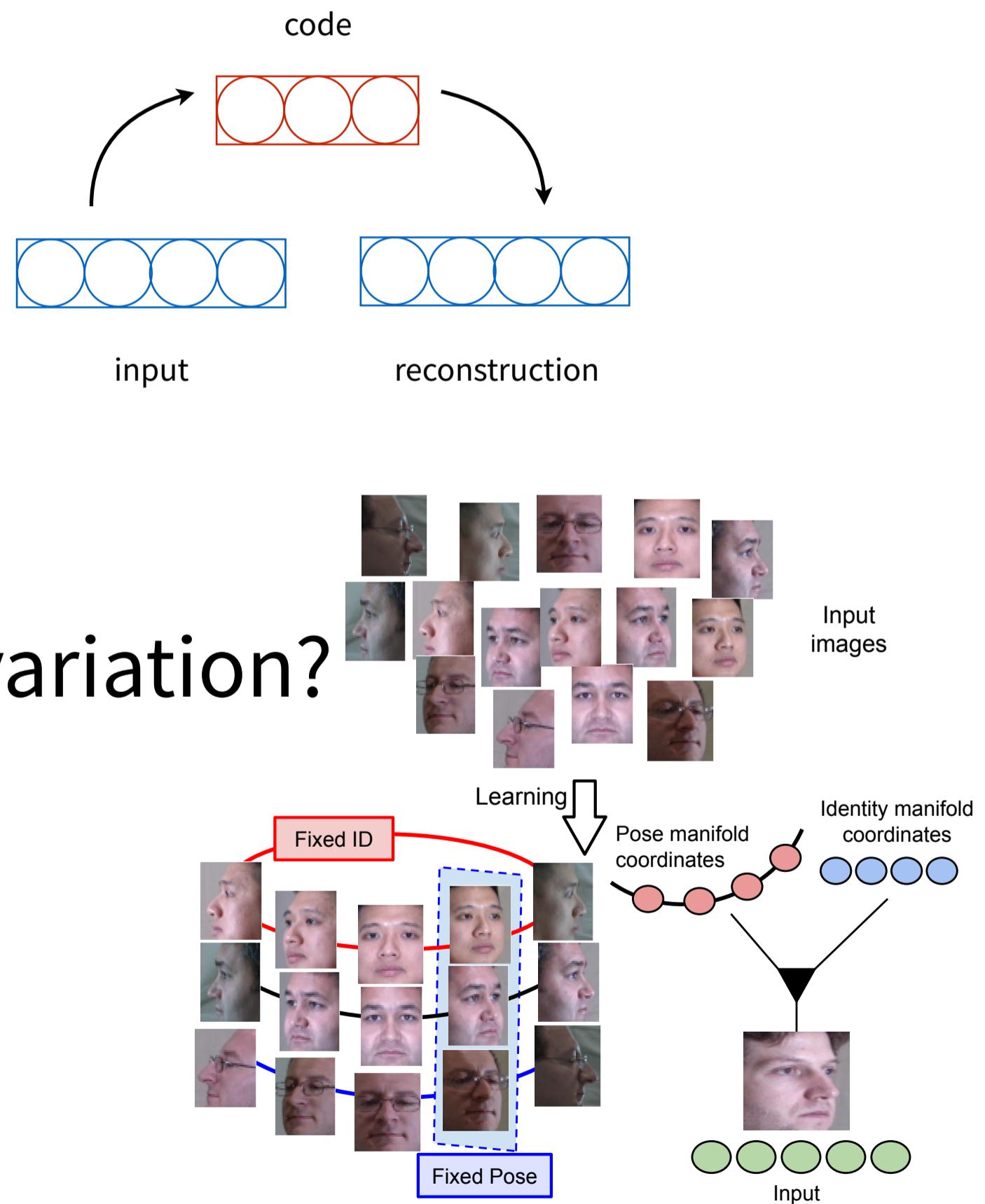


Unsupervised Learning of Representations



Unsupervised learning of representations

- What is the objective?
 - reconstruction error?
 - maximum likelihood?
 - disentangle factors of variation?



Overview (for the remainder of the session)

- Unsupervised building blocks of Deep Learning
 - Auto-encoders
 - Restricted Boltzmann Machines
- Their use in Deep Architectures (how/why?)
- Practical considerations