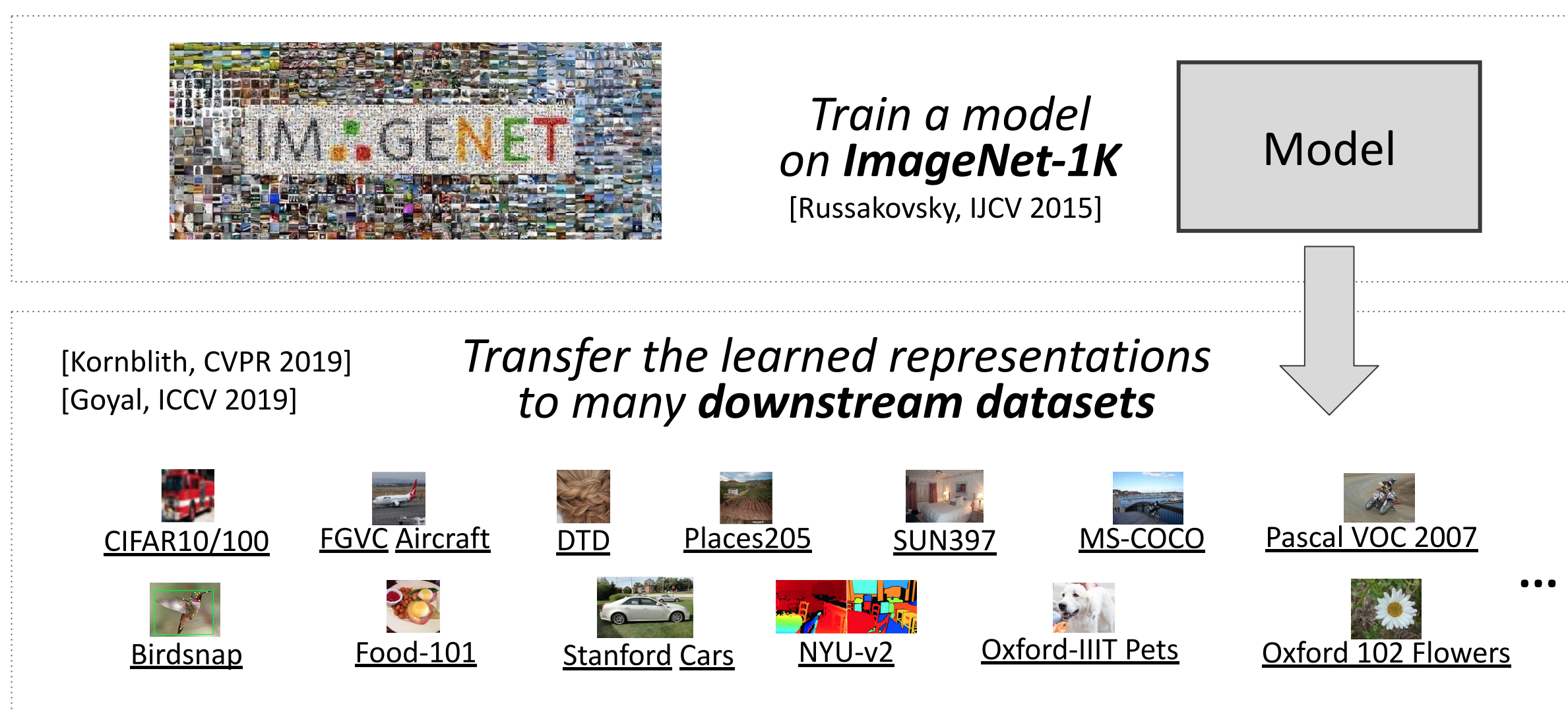


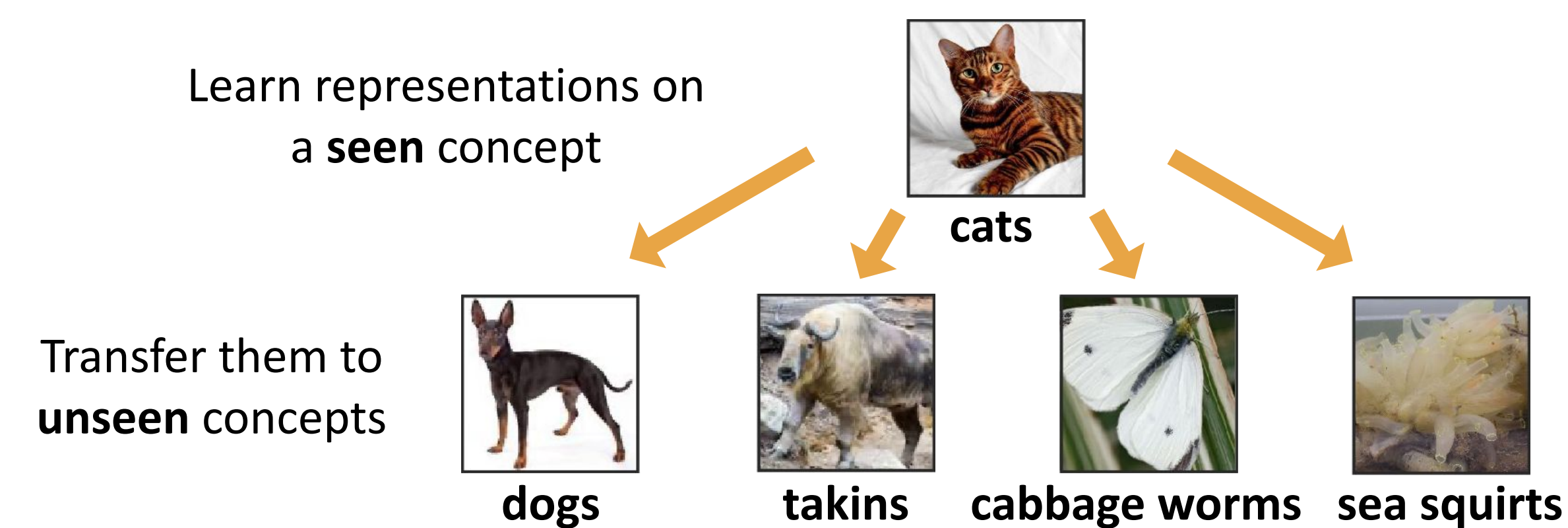
Highlights / Take-home messages

- ➔ We propose the **ImageNet-CoG benchmark**
 - ◆ Enables measuring concept generalization in a principled way
 - ◆ **Seen** concepts ⇒ ImageNet-1K concepts
 - ◆ **Unseen** concepts ⇒ Sampled from the full ImageNet-21K dataset
 - **5 Levels** ⇒ Increasingly challenging transfer datasets
- ➔ To be used **out-of-the-box** for ImageNet-1K pretrained models
- ➔ **31 models evaluated** on ImageNet-CoG
 - ◆ Interesting insights on popular state-of-the-art methods

Learning general-purpose visual representations



Concept generalization



- ➔ No systematic approach for evaluating concept generalization
- ➔ Unknown semantic similarity between ImageNet-1K and other datasets

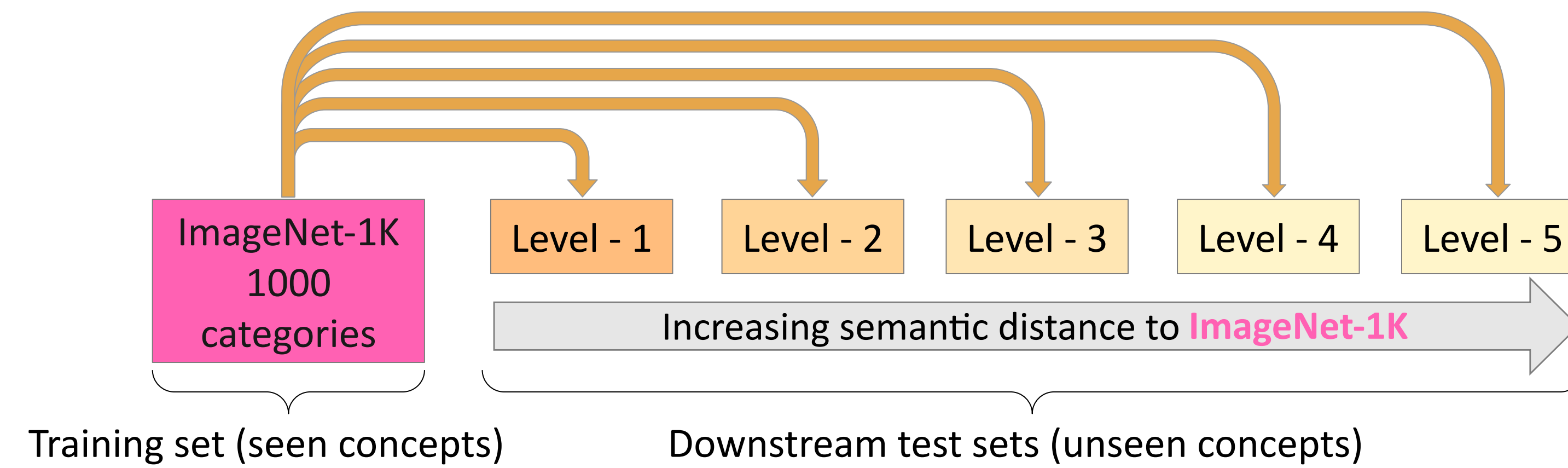
We tackle the following questions:

- How can we evaluate concept generalization reliably?
- Which methods are the best for concept generalization?

Proposed ImageNet-CoG Benchmark

A benchmark tailored for concept generalization, built on the full ImageNet-21K [Deng, CVPR 2009]

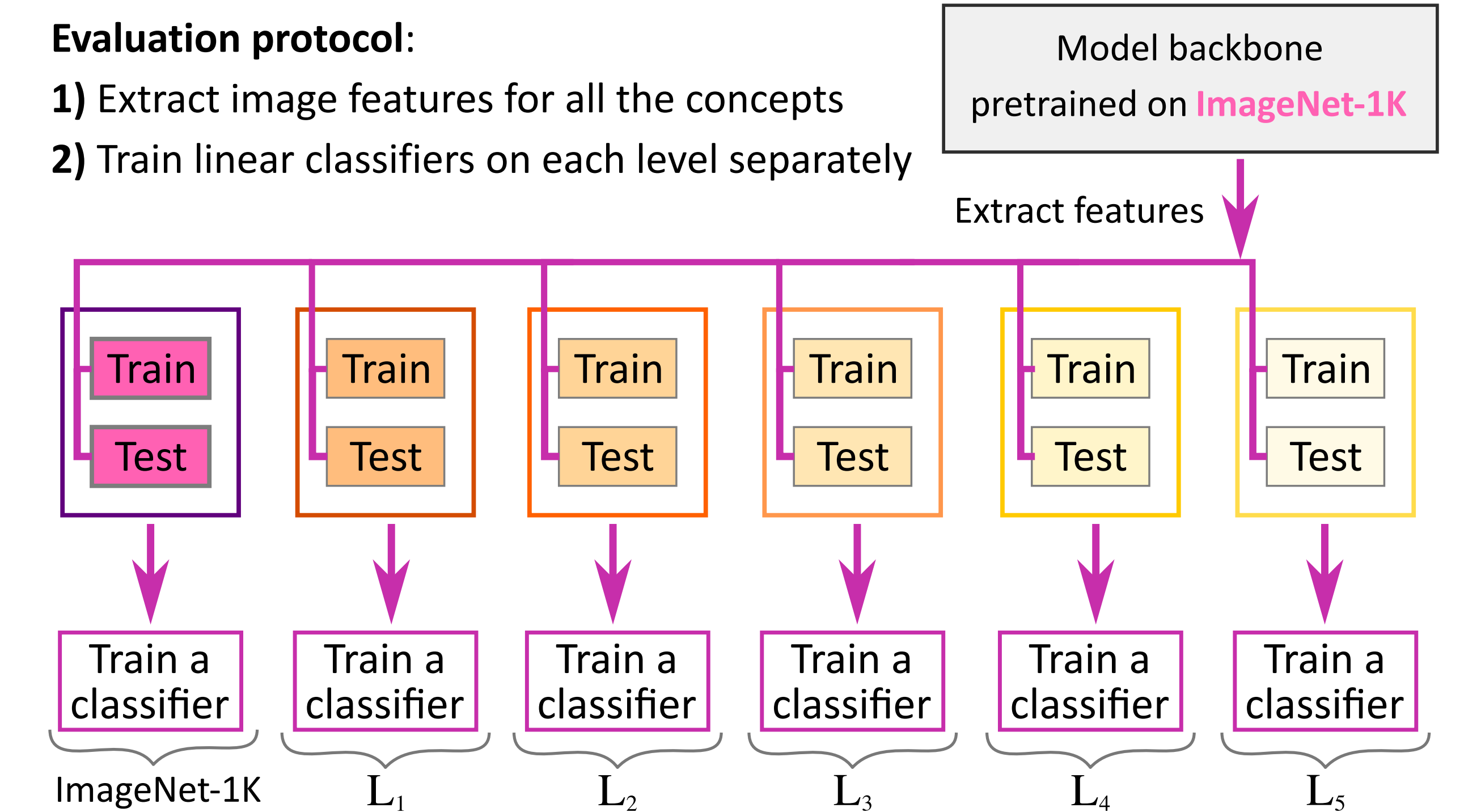
- ➔ Transfer learning scenarios from ImageNet-1K to ImageNet-CoG levels (sampled from the full ImageNet)



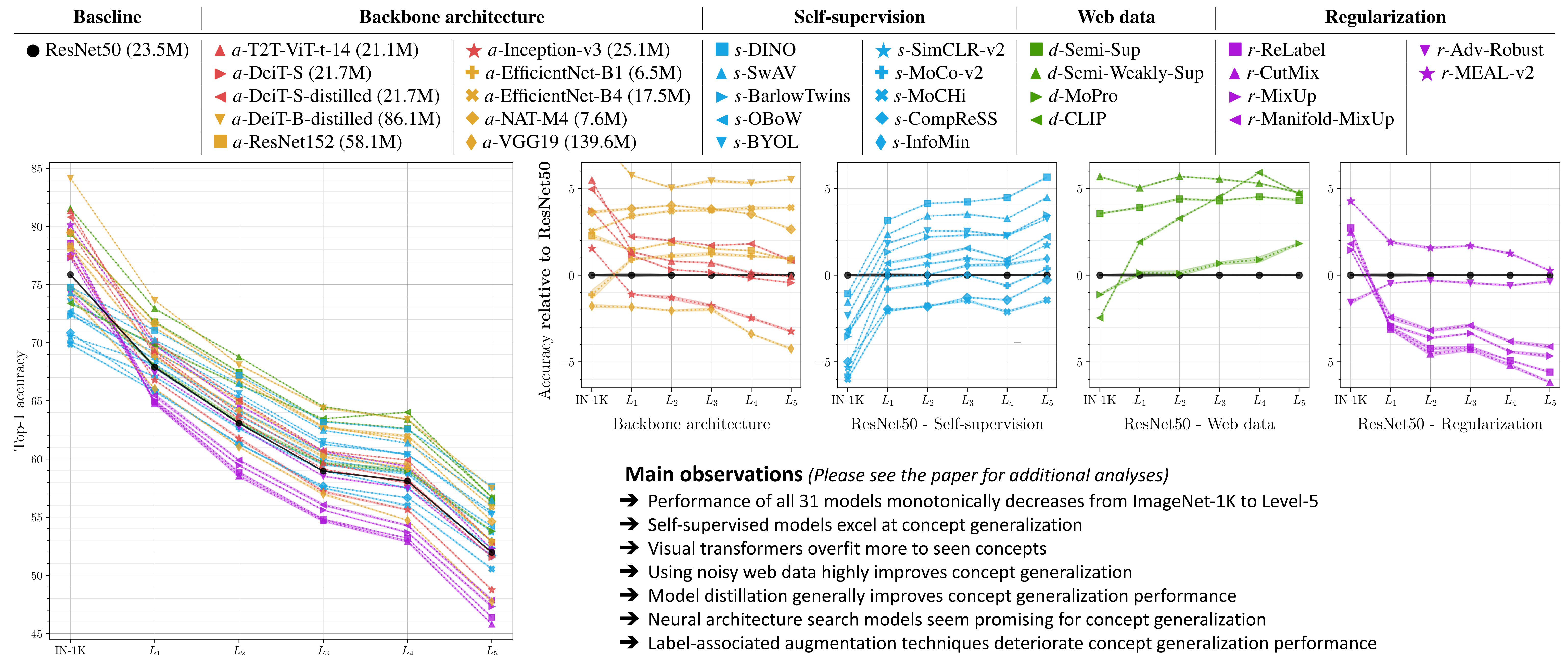
- ➔ Seen and unseen concepts are in the same concept ontology (WordNet ontology [Miller, ACM 1995])
- ➔ Semantic similarity between concepts is defined by linguists

Evaluation protocol:

- 1) Extract image features for all the concepts
- 2) Train linear classifiers on each level separately



Main results of evaluating 31 representation learning methods on the ImageNet-CoG Benchmark



Main observations (Please see the paper for additional analyses)

- ➔ Performance of all 31 models monotonically decreases from ImageNet-1K to Level-5
- ➔ Self-supervised models excel at concept generalization
- ➔ Visual transformers overfit more to seen concepts
- ➔ Using noisy web data highly improves concept generalization
- ➔ Model distillation generally improves concept generalization performance
- ➔ Neural architecture search models seem promising for concept generalization
- ➔ Label-associated augmentation techniques deteriorate concept generalization performance