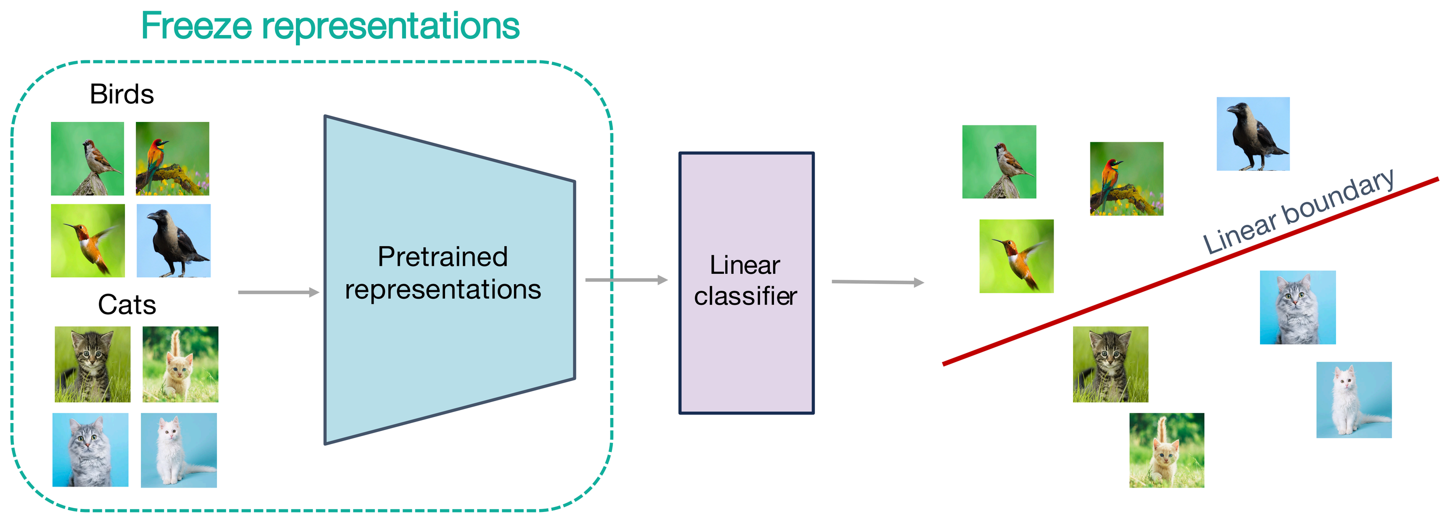
The Pursuit of Human Labeling: A New Perspective on Unsupervised Learning

Artyom Gadetsky and Maria Brbić

EPFL

Motivation

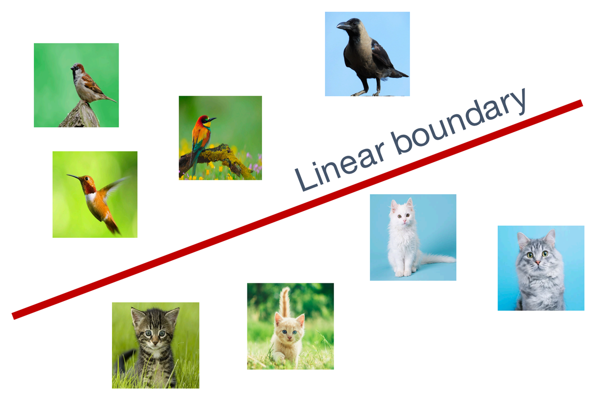
Given pretrained representations, **supervised fine-tuning** is a standard approach to perform transfer learning to solve a new task



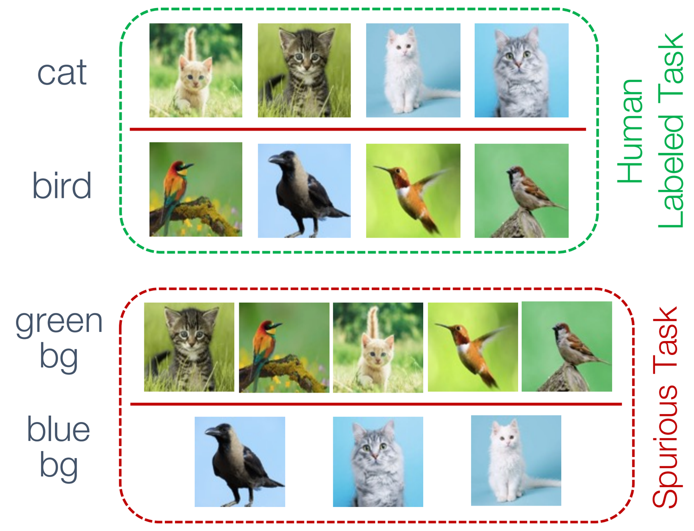
Can we use this paradigm for unsupervised inference of human labeled tasks?

What makes human labeled tasks special?

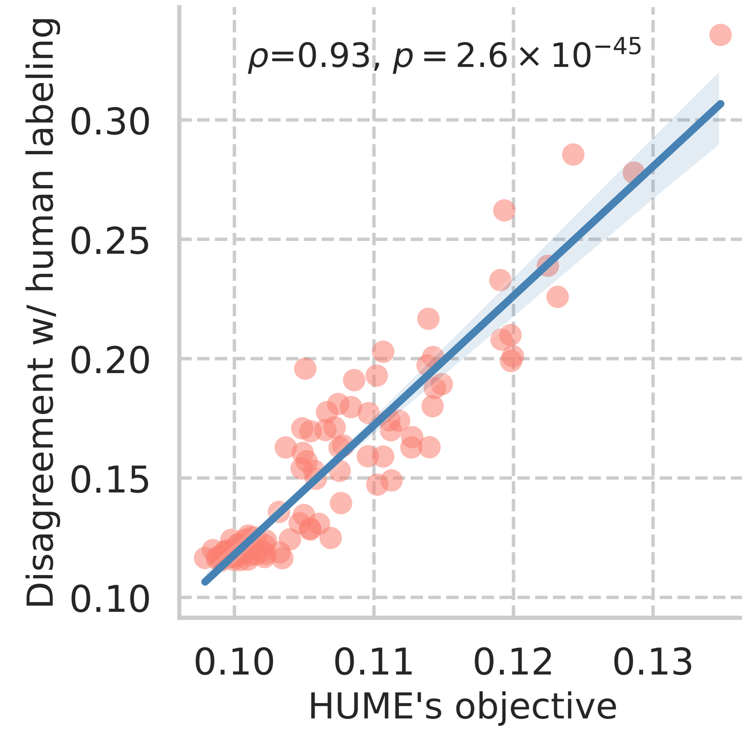
Observation 1:
Many human labeled tasks are linearly separable in a sufficiently strong representation spaces



Observation 2:
Although each representation space has its own inductive biases, human labeled tasks are invariant to the underlying representation space



Results

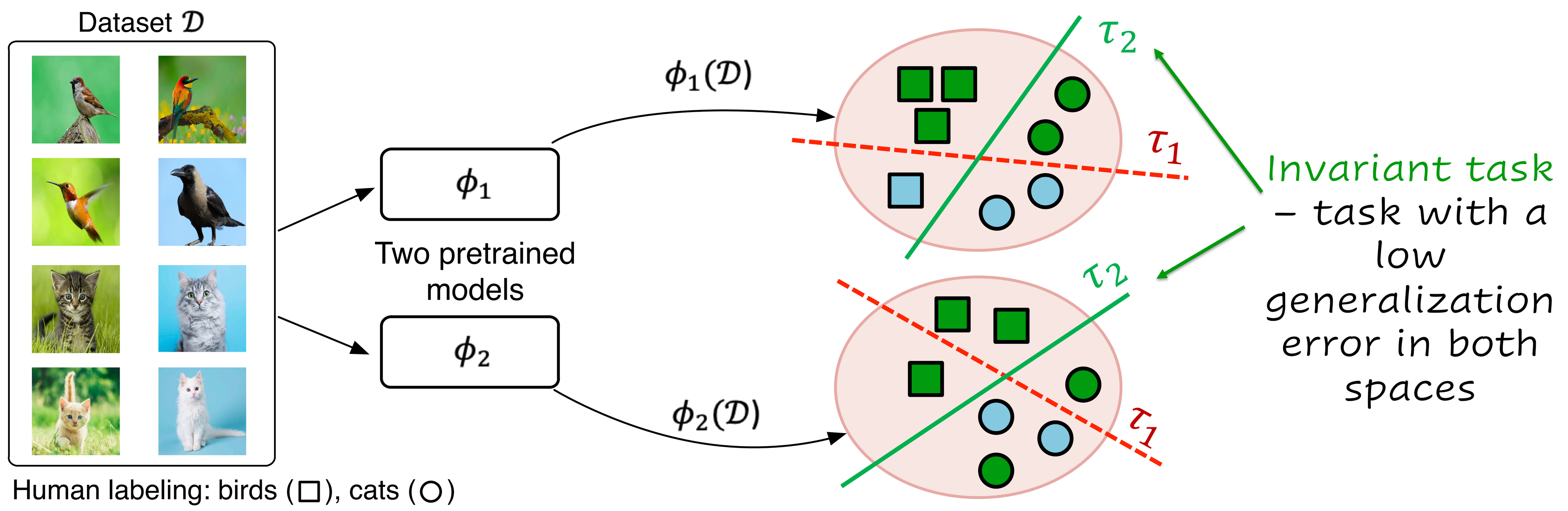


HUME's objective:
Generalization error of linear classifiers in different representation spaces

Agreement with human labeling:
HUME's objective is lower for the tasks that better agree with ground truth human labeled task

HUME trains only linear classifiers on top of pretrained models!

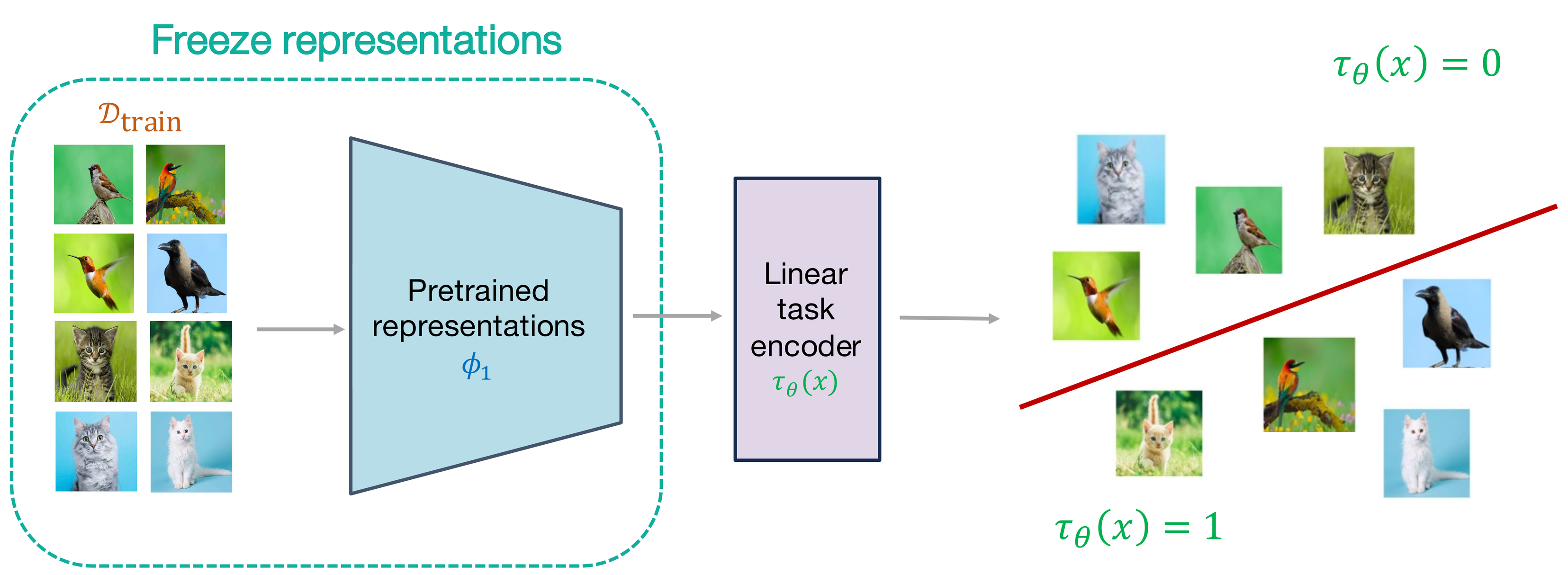
Our approach: HUME



Key idea behind HUME:
Search for the task which attains the lowest generalization error in both representation spaces

Combinatorial optimization problem! 😱
How to solve it in practice?

Step 1:
Label training split $\mathcal{D}_{train}$ using a linear task encoder in the first representation space ϕ_1.



Step 2:
Fit generated labeling on the training split $\mathcal{D}_{train}$ with a linear model in the second representation space ϕ_2:

$$m^*(x) = \arg\min_{m(x):=w^T\phi_2(x)} \mathcal{L}_{\mathcal{D}_{train}}\big(m(x);\ \tau_\theta(x)\big)$$

Step 3:
Minimize generalization error of $m^*(x)$ with respect to a labeling τ_θ on a held-out $\mathcal{D}_{test}$:

$$\min_{\tau_\theta} \mathcal{L}_{\mathcal{D}_{test}}\big(m^*(x);\ \tau_\theta(x)\big)$$

Comparison to supervised fine-tuning:
HUME can match the performance of the supervised model while being fully-unsupervised!

	STL-10		CIFAR-10		CIFAR-100-20	
Method	ACC	ARI	ACC	ARI	ACC	ARI
Supervised FT	88.9	77.7	**89.5**	**79.0**	**72.5**	**52.6**
HUME (Trans.)	**93.2**	**86.0**	89.2	79.2	56.7	39.6

Comparison to unsupervised baselines:
HUME outperforms existing unsupervised baselines by a large margin!

	STL-10		CIFAR-10		CIFAR-100-20	
Method	ACC	ARI	ACC	ARI	ACC	ARI
SCAN	77.8	61.3	83.3	70.5	45.4	29.7
SPICE	86.2	73.2	84.5	70.9	46.8	32.1
HUME (Ind.)	**90.8**	**81.2**	**88.4**	**77.6**	**55.5**	**37.7**

Large-scale unsupervised learning on the ImageNet-1k:

Method	ACC	ARI
SCAN	39.7	27.9
Twist	40.6	30.0
Self-classifier	41.1	29.5
HUME (Ind.)	**51.1**	**38.1**

HUME scales to large datasets and achieves remarkable improvement over existing baselines