ELFS: Label-Free Coreset Selection with Proxy Training Dynamics

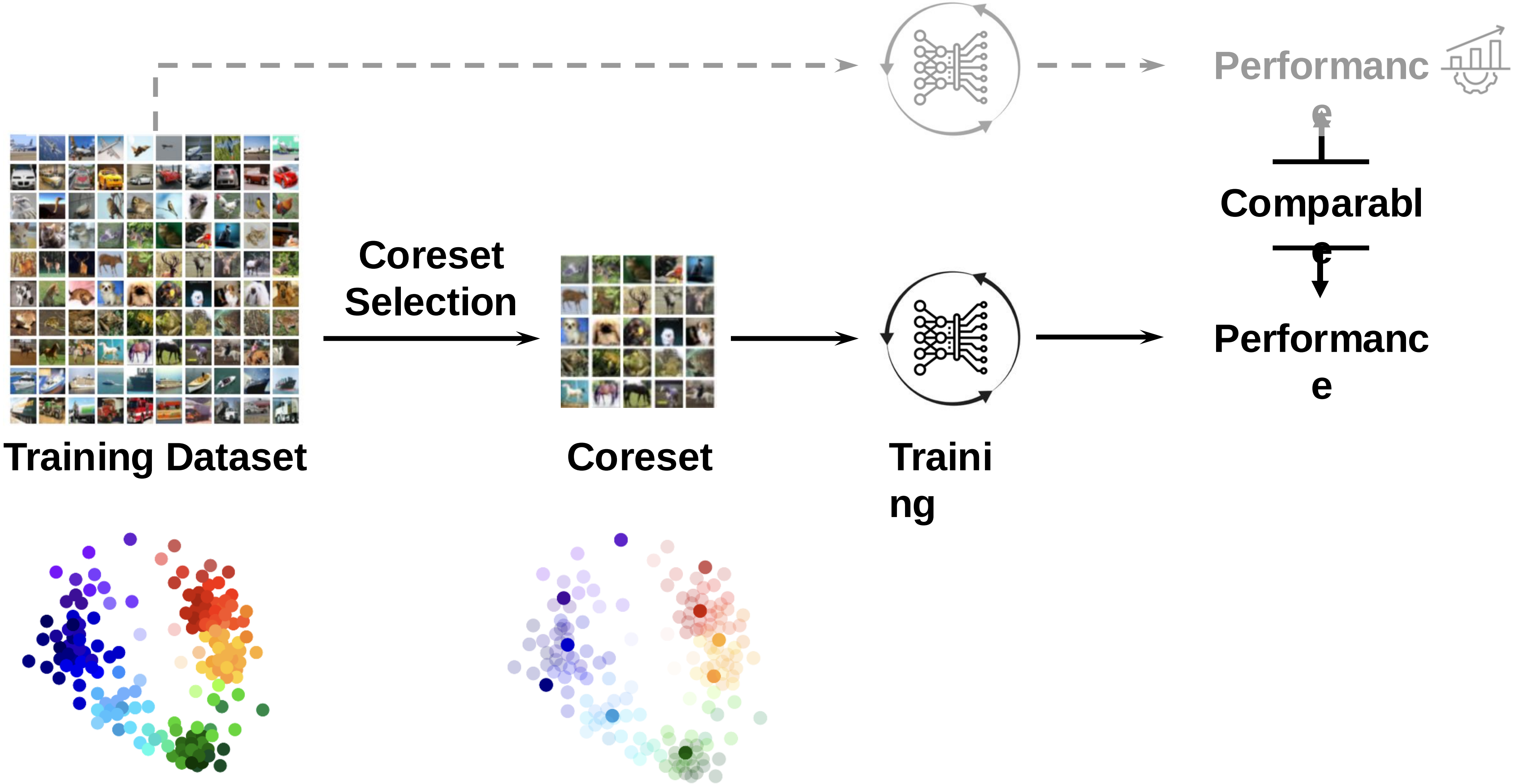
Haizhong Zheng[*1], Elisa Tsai[*1], Yifu Lu[1], Jiachen Sun[1], Brian R. Bartoldson[2], Bhavya Kailkhura[2], Atul Prakash[1]

[1] University of Michigan
[2] Lawrence Livermore National Laboratory

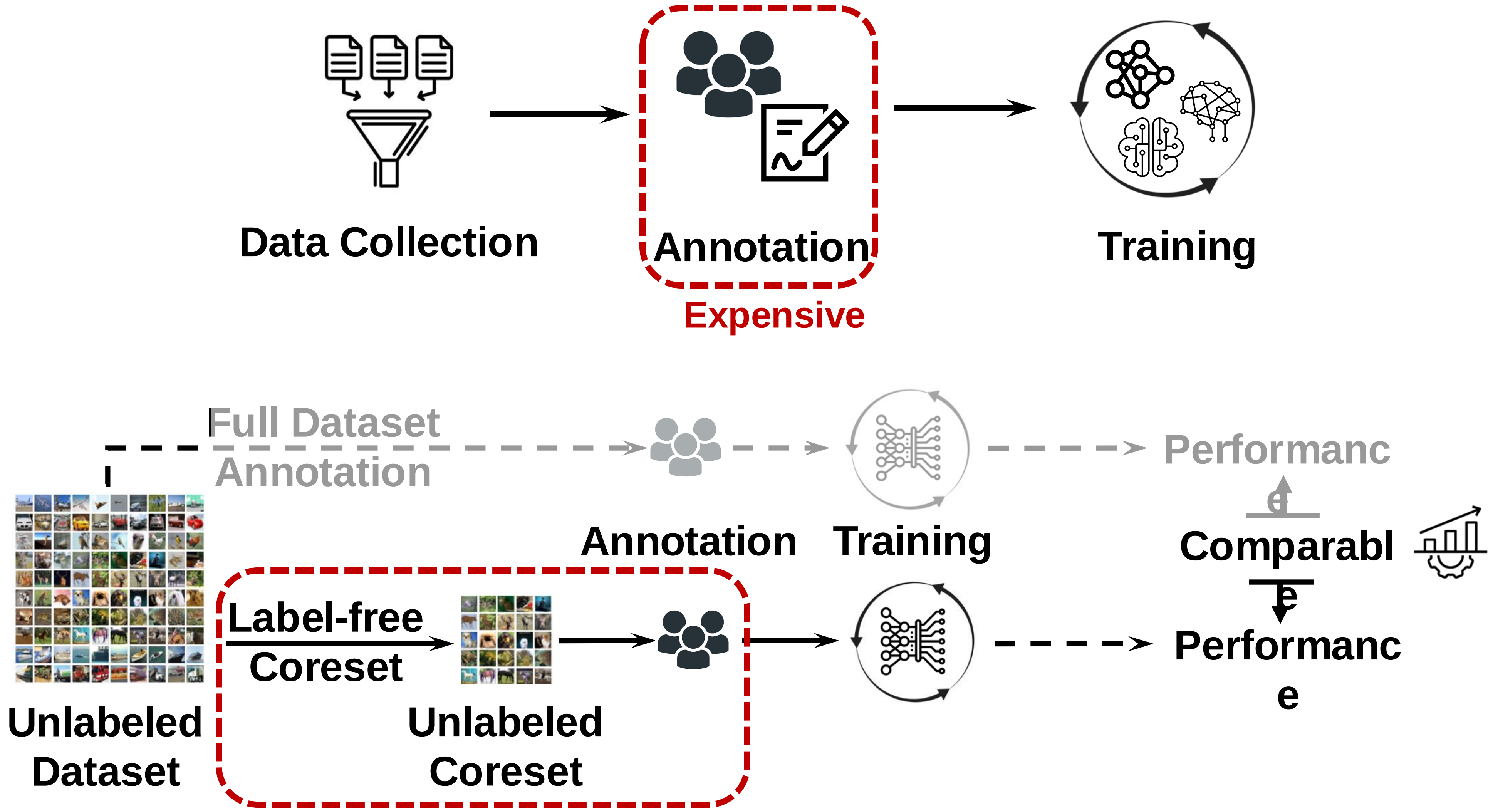
[*] Equal Contribution

UNIVERSITY OF MICHIGAN

Coreset Seletion



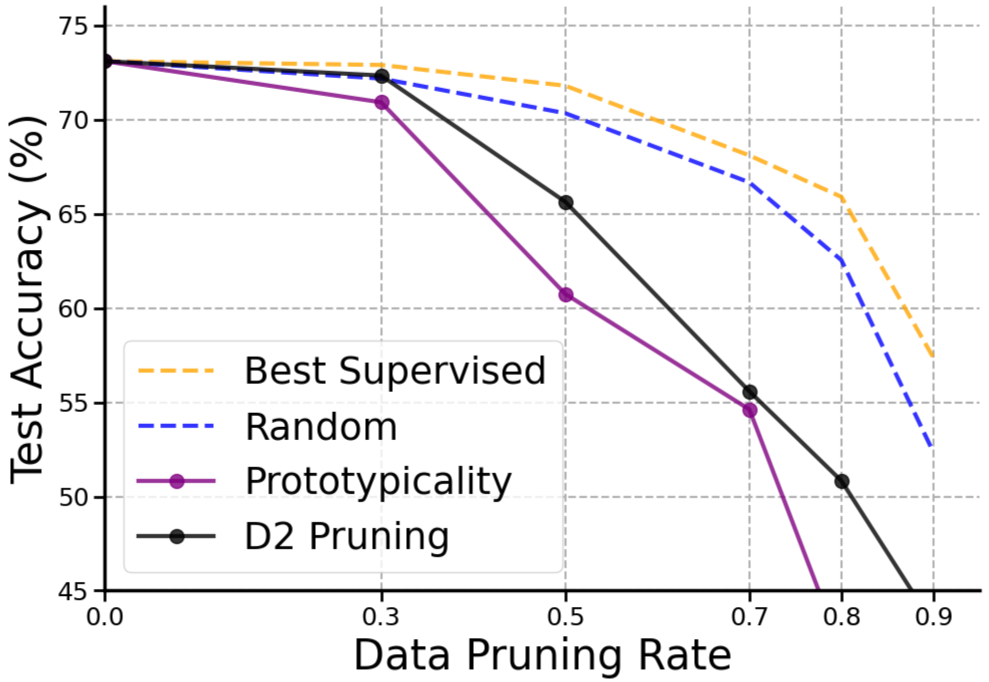
Can Coreset Selection Make Data Collection More Efficient?



Limitations on SOTA Label-free Methods

The dilemma of label-free coreset selection.

Ground-truth labels	→	**Good Performance**
Label-free	→	**Poor Performance**



Label-Free Coreset Selection on ImageNet-1k

Two key challenges:

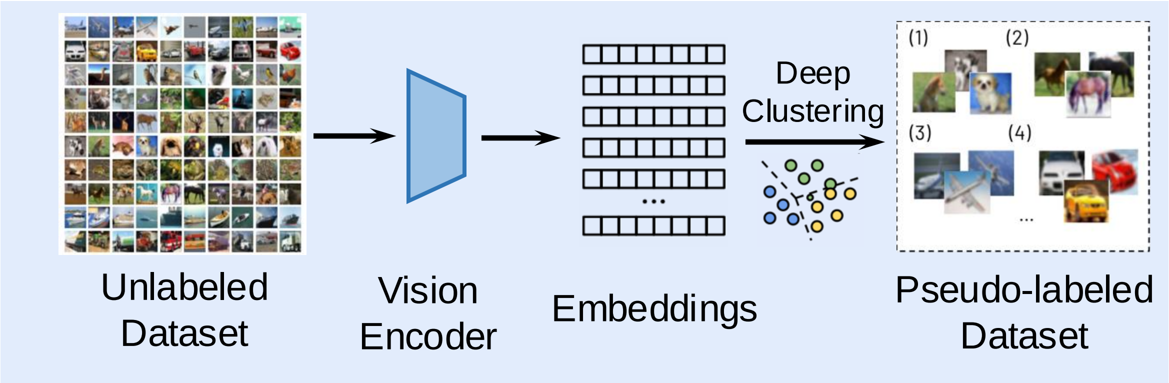
1. **How to estimate training dynamics scores without labels?**
2. **How to select high-quality coresets with proxy scores?**

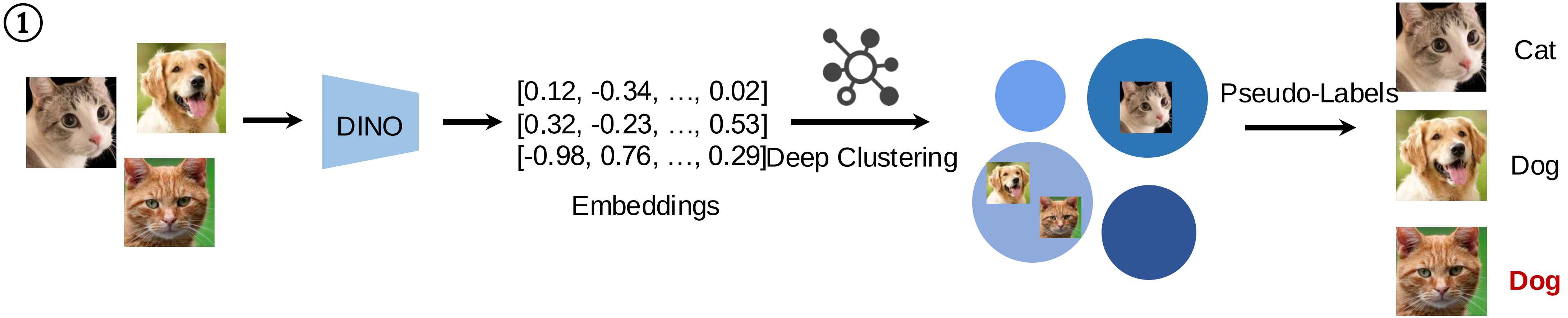
ELFS: Effective Label-Free Coreset Selection

① **Pseudo-Labels Generated by Deep Clustering**

② **Coreset Selection**

③ **Labeling & Training**

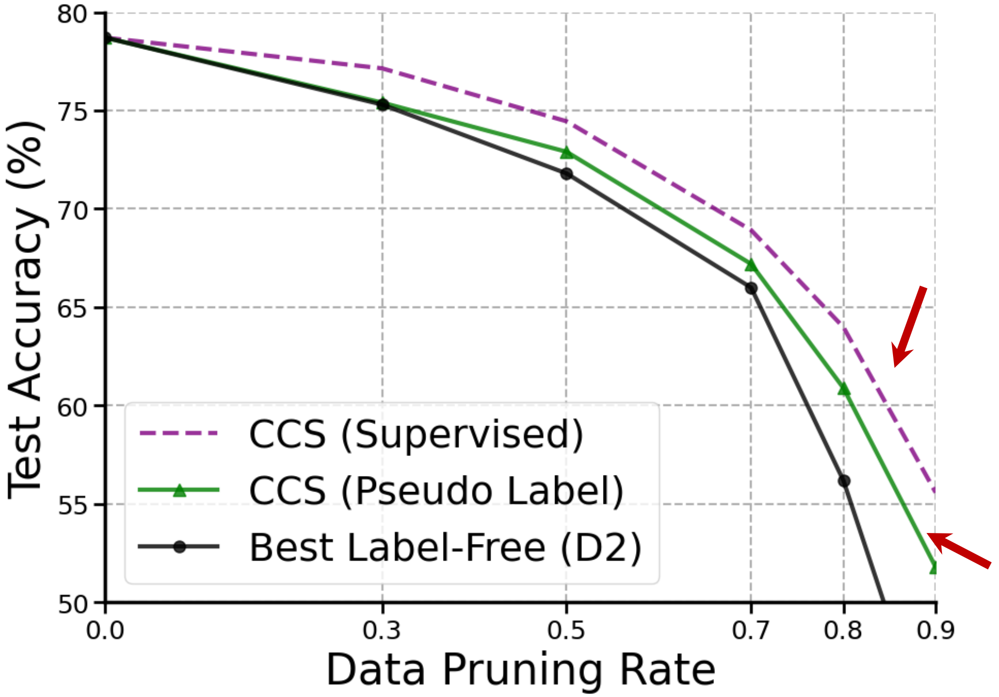


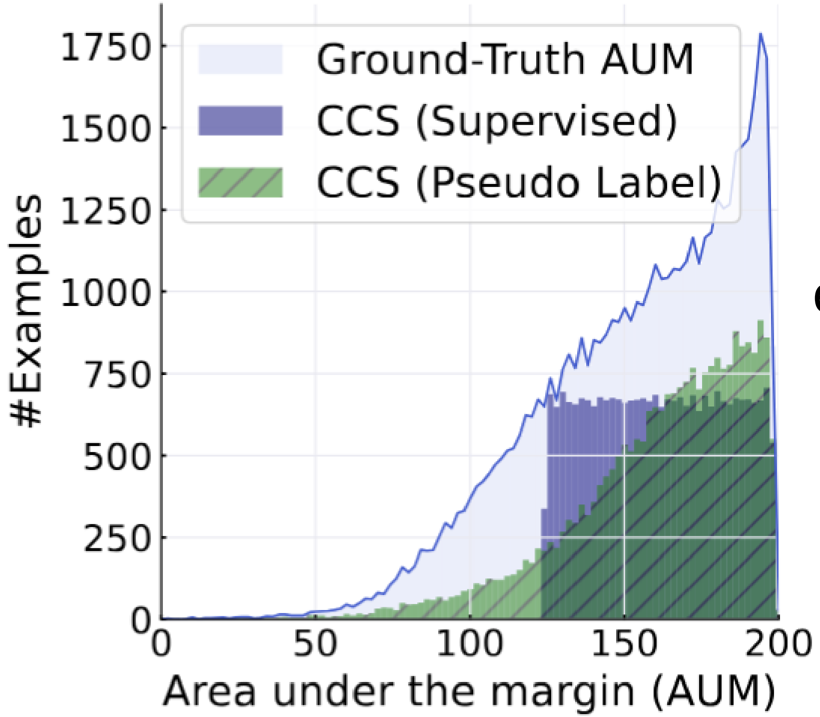


Coreset Selection with Proxy Training Dynamics

Directly apply existing sampling (CCS):

1. Performance gap to supervised coreset selection.
2. Biased score distribution caused by label noises.



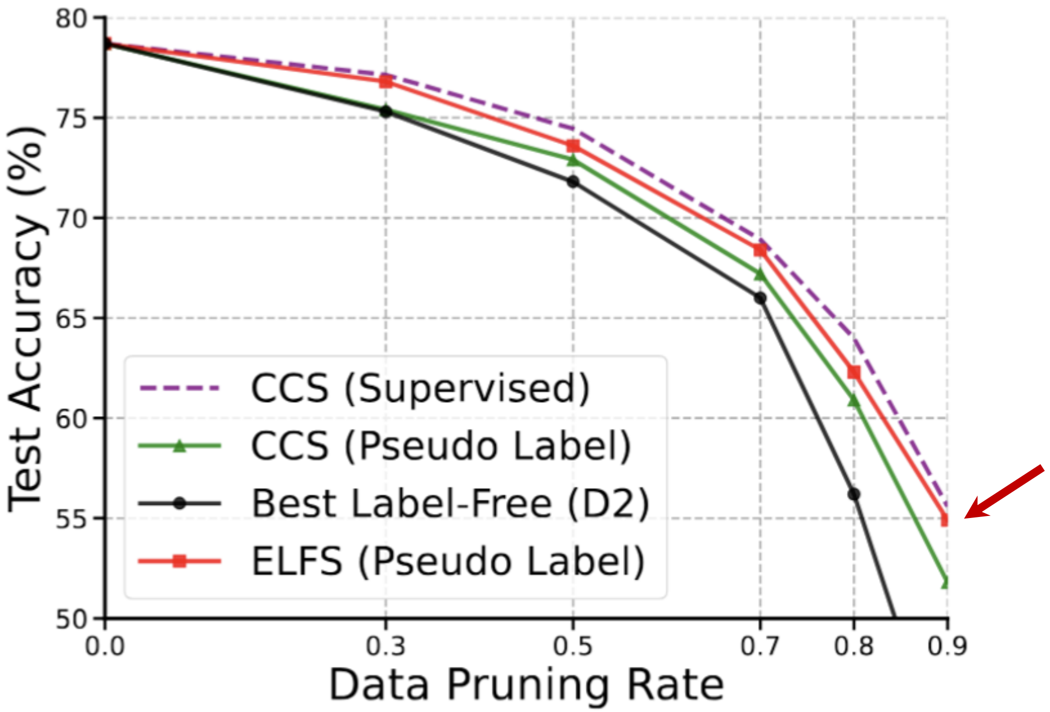


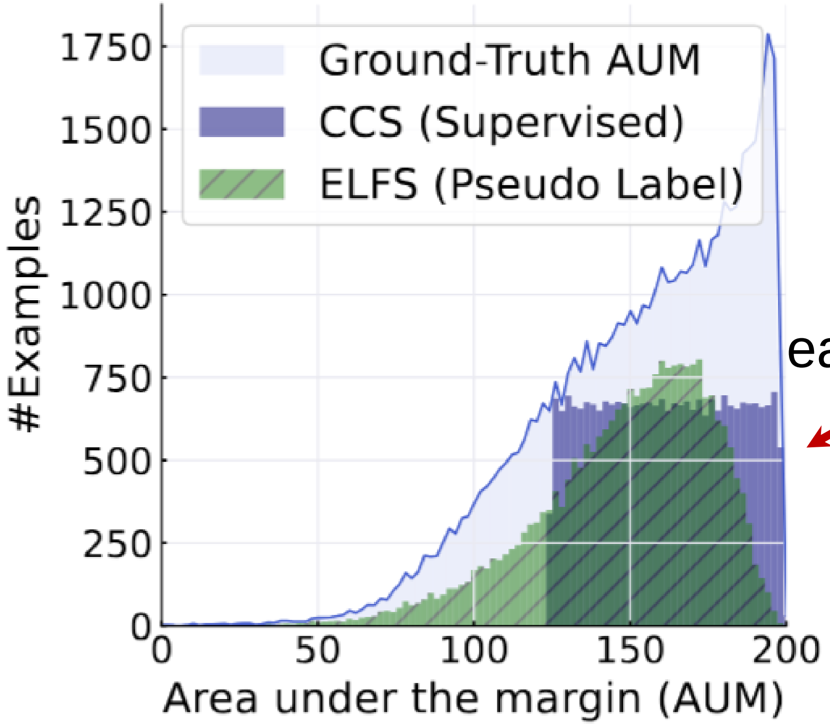
Much more easy examples.

Coreset Selection with Proxy Training Dynamics

Double-End Pruning:

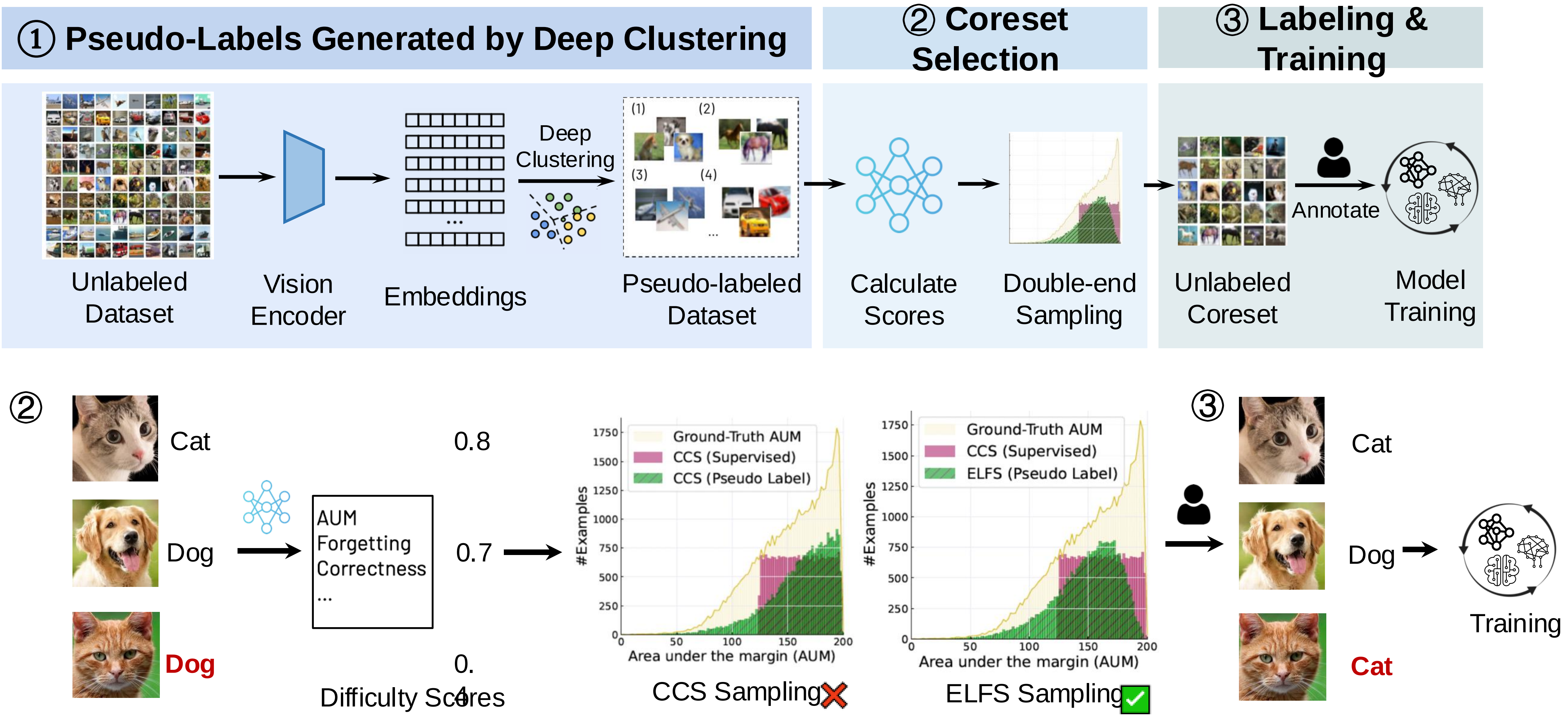
After pruning hard examples, we prune easy examples until the budget is met.





Less easy examples.

ELFS: Effective Label-Free Coreset Selection



Evaluation

Bold: The best with the same encoder. Worse than random.

		CIFAR10					CIFAR100					ImageNet-1K				
Encoder	Pruning Rate	30%	50%	70%	80%	90%	30%	50%	70%	80%	90%	30%	50%	70%	80%	90%
-	Best Supervised	95.7	94.9	93.3	91.4	87.1	78.2	75.9	70.5	65.2	56.9	72.9	71.8	68.1	65.9	55.6
-	Random	94.3	93.4	90.9	88.0	79.0	74.6	71.1	65.3	57.4	44.8	72.2	70.3	66.7	62.5	52.3
	Badge (AL)	93.6	93.0	91.0	87.9	81.6	74.7	71.8	65.2	58.9	47.8	71.7	70.4	65.8	61.7	53.4
SwAV	Prototypicality	94.7	92.9	90.1	84.2	70.9	74.5	69.8	61.1	48.3	32.1	70.9	60.8	54.6	41.9	30.6
	D2	94.3	93.8	91.6	85.1	71.4	75.3	71.3	**66.0**	56.2	42.1	72.3	65.6	55.6	50.8	43.2
	ELFS (Ours)	**95.0**	**94.3**	**91.8**	**89.8**	**82.5**	**76.1**	**72.1**	65.5	**58.2**	**49.8**	**73.2**	**71.4**	**66.8**	**62.7**	**53.4**
DINO	FreeSel	94.5	93.8	91.7	88.9	82.4	75.0	70.5	65.6	57.6	44.8	72.2	70.0	65.4	61.0	51.1
	ELFS (Ours)	**95.5**	**95.2**	**93.2**	**90.7**	**87.3**	**76.8**	**73.6**	**68.4**	**62.3**	**54.9**	**73.5**	**71.8**	**67.2**	**63.4**	**54.9**

1. ELFS consistently outperform other baselines.

Evaluation

Bold: The best with the same encoder. Worse than random.

		CIFAR10					CIFAR100					ImageNet-1K				
Encoder	Pruning Rate	30%	50%	70%	80%	90%	30%	50%	70%	80%	90%	30%	50%	70%	80%	90%
-	Best Supervised	95.7	94.9	93.3	91.4	87.1	78.2	75.9	70.5	65.2	56.9	72.9	71.8	68.1	65.9	55.6
-	Random	94.3	93.4	90.9	88.0	79.0	74.6	71.1	65.3	57.4	44.8	72.2	70.3	66.7	62.5	52.3
	Badge (AL)	93.6	93.0	91.0	87.9	81.6	74.7	71.8	65.2	58.9	47.8	71.7	70.4	65.8	61.7	53.4
SwAV	Prototypicality	94.7	92.9	90.1	84.2	70.9	74.5	69.8	61.1	48.3	32.1	70.9	60.8	54.6	41.9	30.6
	D2	94.3	93.8	91.6	85.1	71.4	75.3	71.3	**66.0**	56.2	42.1	72.3	65.6	55.6	50.8	43.2
	ELFS (Ours)	**95.0**	**94.3**	**91.8**	**89.8**	**82.5**	**76.1**	**72.1**	65.5	**58.2**	**49.8**	**73.2**	**71.4**	**66.8**	**62.7**	**53.4**
DINO	FreeSel	94.5	93.8	91.7	88.9	82.4	75.0	70.5	65.6	57.6	44.8	72.2	70.0	65.4	61.0	51.1
	ELFS (Ours)	**95.5**	**95.2**	**93.2**	**90.7**	**87.3**	**76.8**	**73.6**	**68.4**	**62.3**	**54.9**	**73.5**	**71.8**	**67.2**	**63.4**	**54.9**

1. ELFS consistently outperform other baselines.
2. ELFS matches supervised performance for some pruning rates.

Evaluation

Bold: The best with the same encoder. Worse than random.

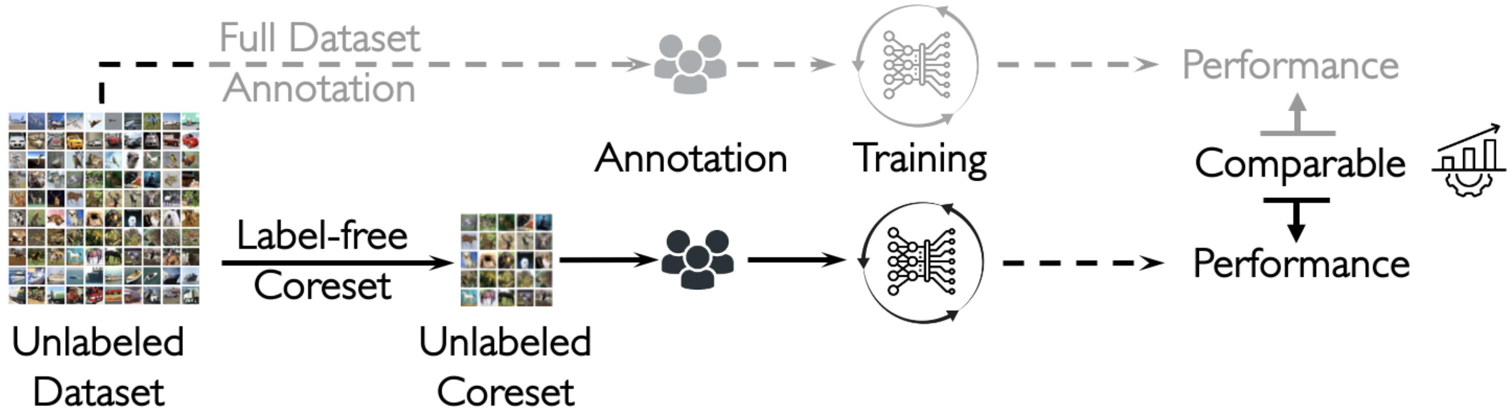
		CIFAR10					CIFAR100					ImageNet-1K				
Encoder	Pruning Rate	30%	50%	70%	80%	90%	30%	50%	70%	80%	90%	30%	50%	70%	80%	90%
-	Best Supervised	95.7	94.9	93.3	91.4	87.1	78.2	75.9	70.5	65.2	56.9	72.9	71.8	68.1	65.9	55.6
-	Random	94.3	93.4	90.9	88.0	79.0	74.6	71.1	65.3	57.4	44.8	72.2	70.3	66.7	62.5	52.3
	Badge (AL)	93.6	93.0	91.0	87.9	81.6	74.7	71.8	65.2	58.9	47.8	71.7	70.4	65.8	61.7	53.4
SwAV	Prototypicality	94.7	92.9	90.1	84.2	70.9	74.5	69.8	61.1	48.3	32.1	70.9	60.8	54.6	41.9	30.6
	D2	94.3	93.8	91.6	85.1	71.4	75.3	71.3	**66.0**	56.2	42.1	72.3	65.6	55.6	50.8	43.2
	ELFS (Ours)	**95.0**	**94.3**	**91.8**	**89.8**	**82.5**	**76.1**	**72.1**	65.5	**58.2**	**49.8**	**73.2**	**71.4**	**66.8**	**62.7**	**53.4**
DINO	FreeSel	94.5	93.8	91.7	88.9	82.4	75.0	70.5	65.6	57.6	44.8	72.2	70.0	65.4	61.0	51.1
	ELFS (Ours)	**95.5**	**95.2**	**93.2**	**90.7**	**87.3**	**76.8**	**73.6**	**68.4**	**62.3**	**54.9**	**73.5**	**71.8**	**67.2**	**63.4**	**54.9**

1. ELFS consistently outperform other baselines.
2. ELFS matches supervised performance for some pruning rates.
3. ELFS shows robustn

Pseudo-label Accuracy →

	CIFAR10	CIFAR100	ImageNet-1K
TEMI (SwAV)	60.7%	39.8%	43.1%
TEMI (DINO)	92.5%	66.3%	58.8%

Summary



Contribution:

1. We present ELFS, a novel label-free coreset selection method
   1. Approximate data difficulty scores with pseudo-labels generated by deep clustering
   2. Double-end pruning to select a coreset with inaccurate data difficulty scores.

2. ELFS consistently outperforms SOTA label-free coreset selection baselines.

UNIVERSITY OF MICHIGAN