



universität
wien

mcm

Munich Center for Machine Learning

ECML
PKDD
2024

FALCUN: A Simple and Efficient Deep Active Learning Strategy

Sandra Gilhuber^{1,2}, Anna Beer³, Yunpu Ma¹, and Thomas Seidl^{1,2}

¹Ludwig-Maximilians-Universität München, Munich, Germany

²Munich Center for Machine Learning, Munich, Germany

³ University of Vienna, Vienna, Austria

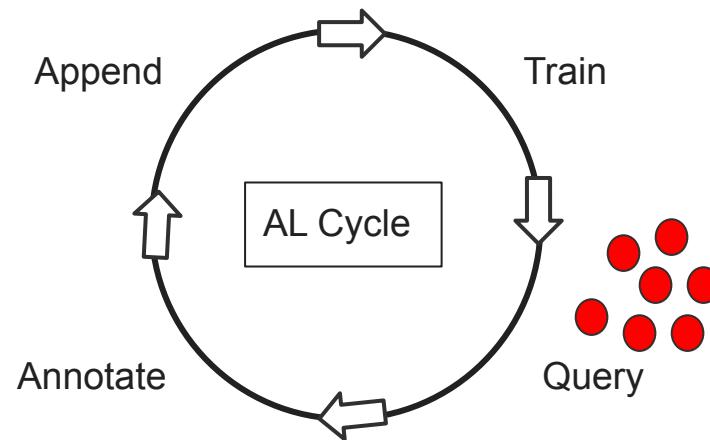
ECML PKDD 2024, Vilnius



<https://github.com/sobermeier/falcun>

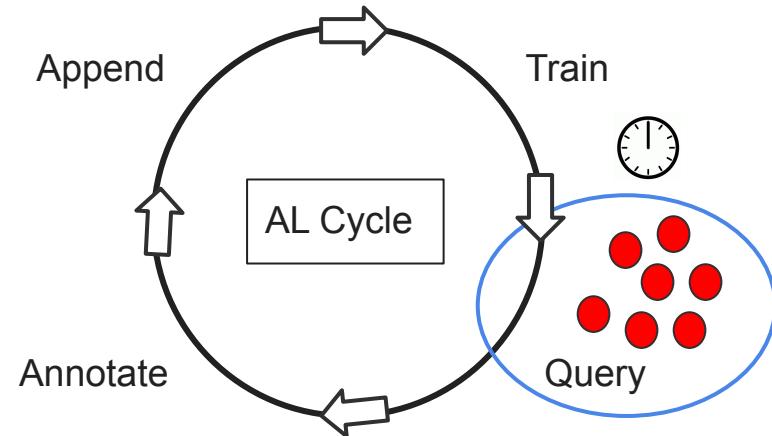
Background

- Active Learning → Batch Active Learning → Deep Batch Active Learning
- Diversity for less redundancy
- Uncertainty for high label-efficiency



Motivation

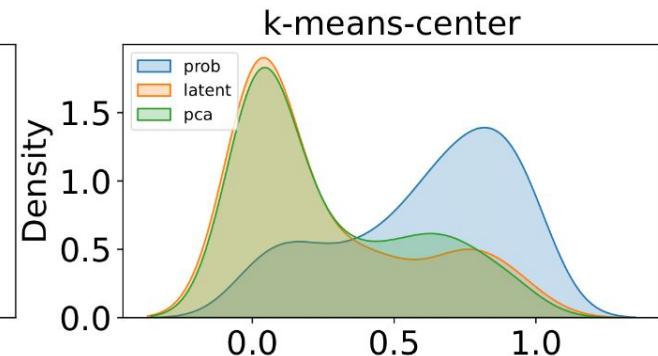
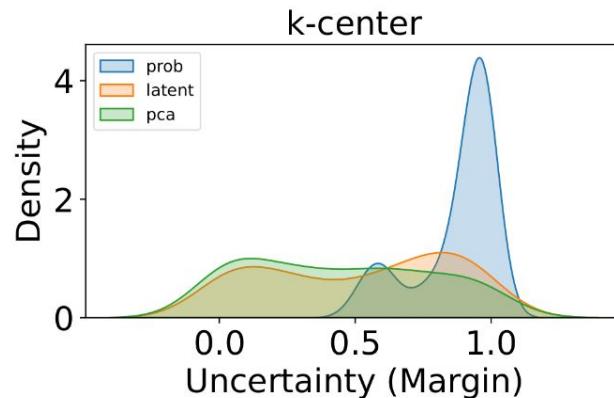
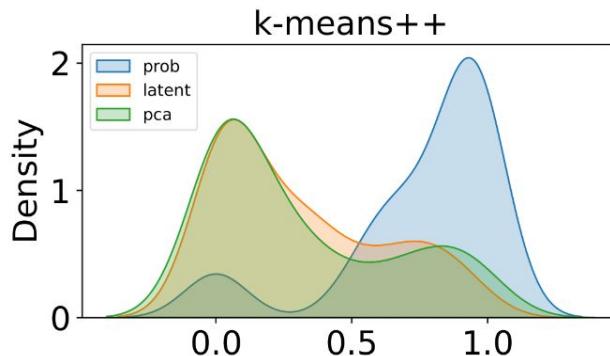
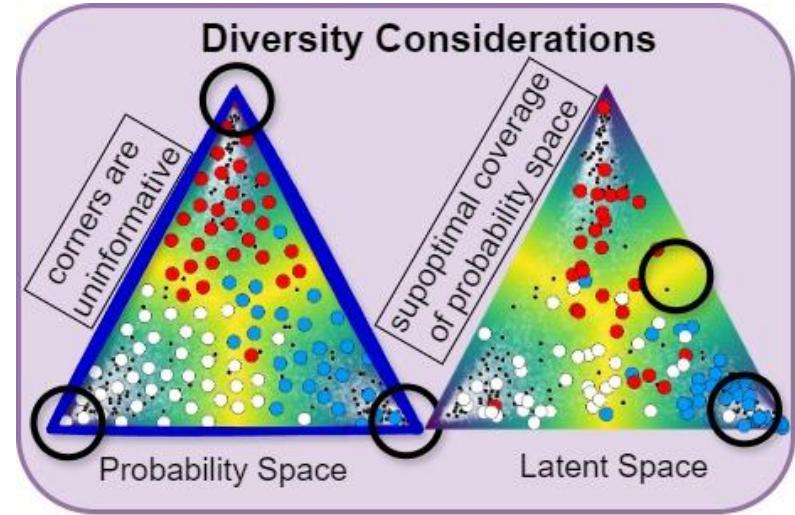
- Sometimes, the *query time* is relevant, too:
 - Domain experts for complex annotation tasks might be available for a limited time window
- Latent features are usually very high-dimensional
→ state-of-the-art methods do not scale well
- Combining diversity and uncertainty
 - is often complex,
 - requires hyperparameters,
 - or has suboptimal trade-offs



Our Approach: FALCUN

Observation 1:

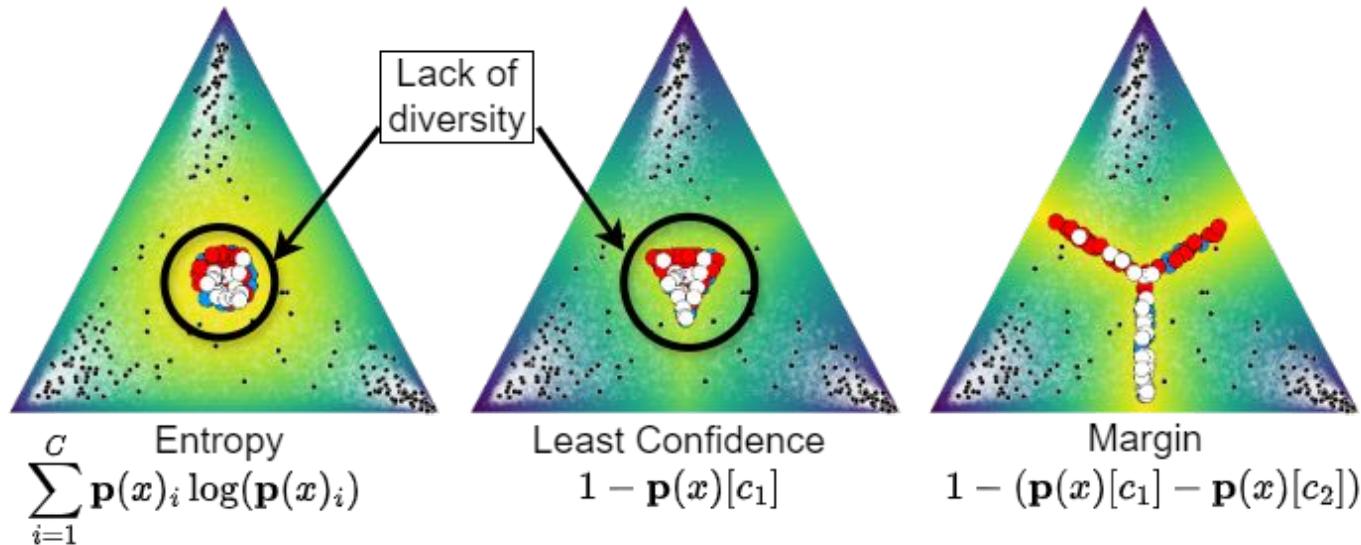
Diversity sampling in the probability space
selects uncertain instances



Our Approach: FALCUN

Observation 2:

Margin uncertainty emphasizes more **diverse** regions than other simple heuristics



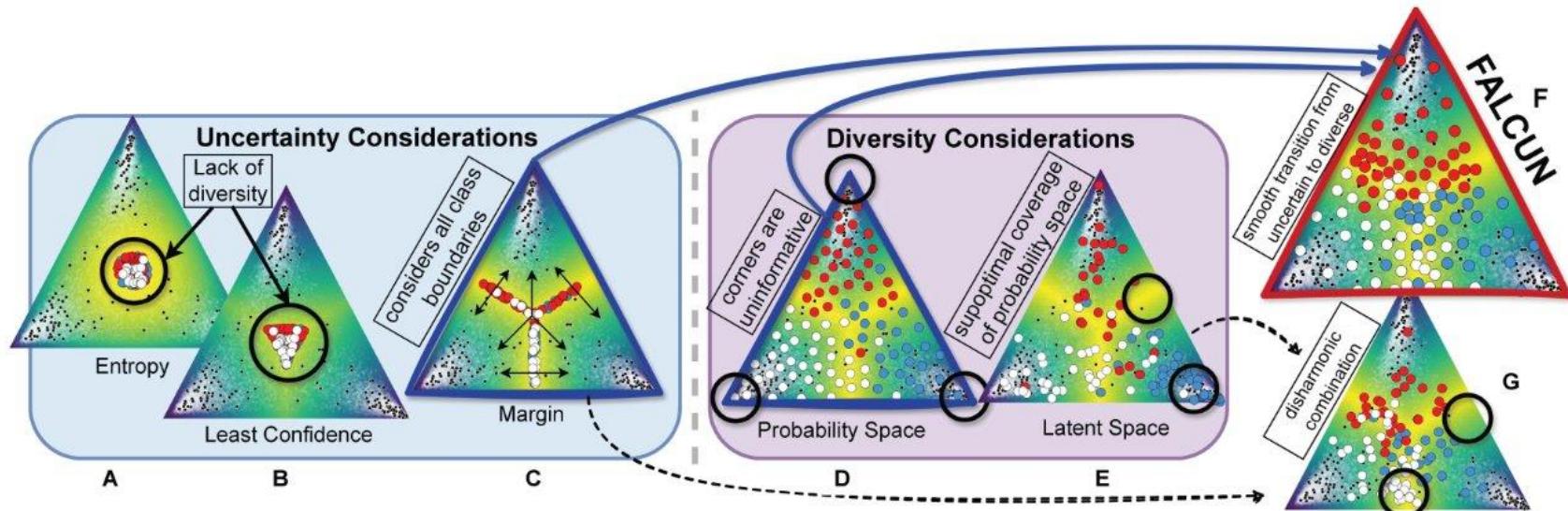
→ Combine diversity $d(x)$ and uncertainty $u(x)$ into one holistic relevance score $r(x)$

Our Approach: FALCUN

Observation 1: **Diversity sampling in the probability space** selects uncertain instances

Observation 2: **Margin uncertainty** emphasizes more **diverse** regions than other simple heuristics

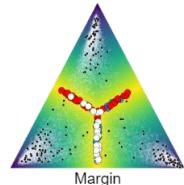
→ Combine diversity $d(x)$ and uncertainty $u(x)$ into one holistic relevance score $r(x)$



Our Approach: FALCUN

- Calculate margin uncertainty $u(x) := 1 - (\mathbf{p}(x)[c_1] - \mathbf{p}(x)[c_2])$
- Initialize diversity score: $d_{init}(x) := u(x)$
- Repeat until query budget is empty:
 - Calculate relevance score $r(x) := u(x) + d(x)$
 - Select query instance x_q with probability proportional to

$$x_q \sim \frac{r(x)^\gamma}{\sum_{x \in \mathcal{U}} r(x)^\gamma}$$



- Update diversity scores: $d(x) \leftarrow \min(d(x), dist(\mathbf{p}(x), \mathbf{p}(x_q)))$
- Normalize diversity values to $[0, 1]$ to align it with uncertainty scores

Advantages

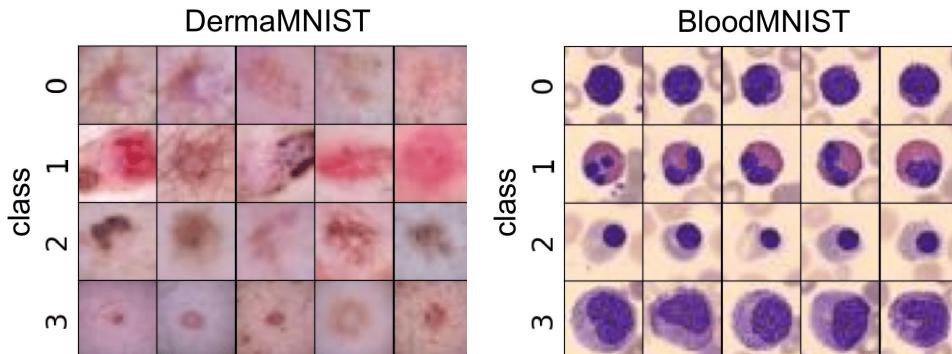
- Faster than competitors that account for diversity within a batch
- No parameters to trade-off between uncertainty and diversity
- Gradually shifts focus from uncertain to more diverse regions
- Easy to understand and implement

Experiments

Table 1: Data set properties: number of points N , number of classes C , and number of input features F .

Type	Data set	N	C	F
Image (Gray)	MNIST	60,000	10	28x28
	RMNIST	60,000	10	28x28
	FashionMNIST	60,000	10	28x28
	EMNIST	131,600	47	28x28
Image (Color)	SVHN	73,257	10	32x32x3
	BloodMNIST	11,959	8	28x28x3
	DermaMNIST	7,007	7	28x28x3
	CIFAR10	60,000	10	28x28x3
Tabular	OpenML-6	16,000	26	17
	OpenML-156	800,000	5	11
	OpenML-155	829,201	10	11

- Model Architectures: MLP, LeNet, Resnet18, Resnet50
- with and without pretraining,
- continual, from scratch



Results - Dueling Matrix shows Label Efficiency

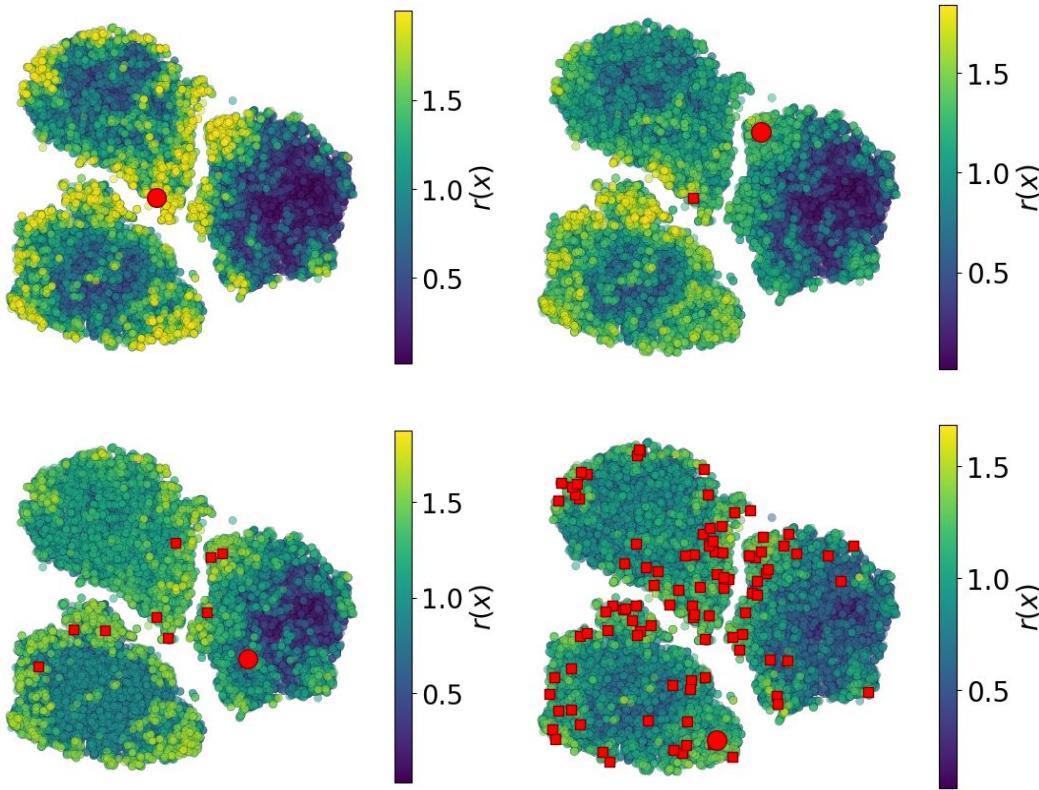
	FALCUN	Badge	AlfaMix	CDAL	CLUE	KCenterGreedy	Entropy	Random	Average Wins (%)
FALCUN	0	10	17	31	37	54	51	71	34
Badge	5	0	17	25	33	48	47	69	30
AlfaMix	3	4	0	24	30	48	51	58	27
CDAL	1	1	11	0	27	39	41	53	22
CLUE	5	3	13	5	0	21	28	46	15
KCenterGreedy	1	2	10	9	3	0	29	39	12
Entropy	0	1	0	1	14	24	0	34	9
Random	0	1	3	13	2	16	31	0	8
Average Losses (%)	2	3	9	13	18	31	35	46	

Row: wins, Column: losses

FALCUN:

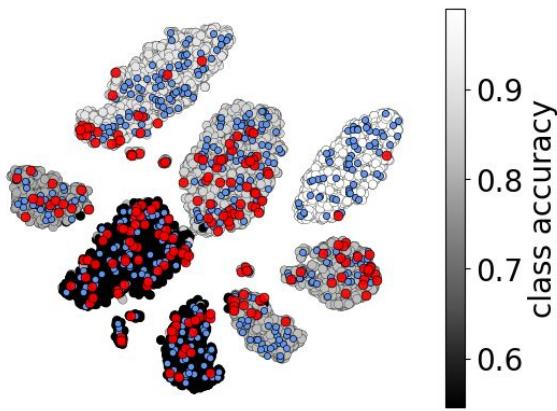
- Most average wins
- Fewest average losses
- Most wins over Random
- No losses against Random

Results - Qualitative



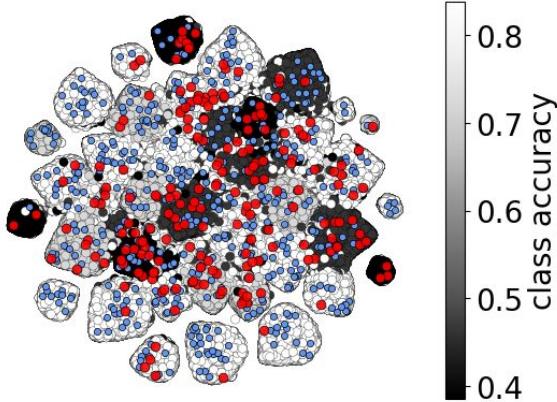
Gradual shift from most uncertain to diverse regions

Results - Qualitative

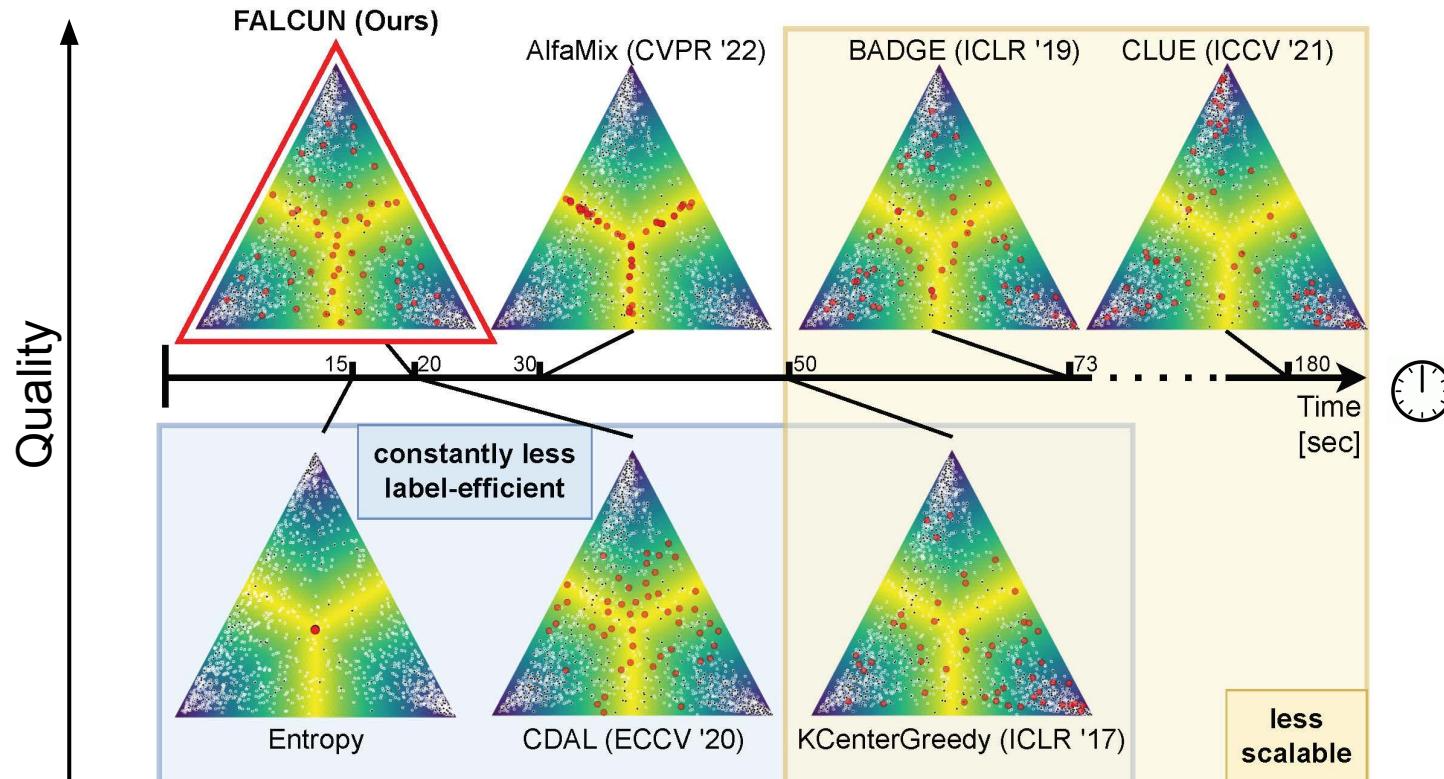


Hue indicates accuracy for a specific class

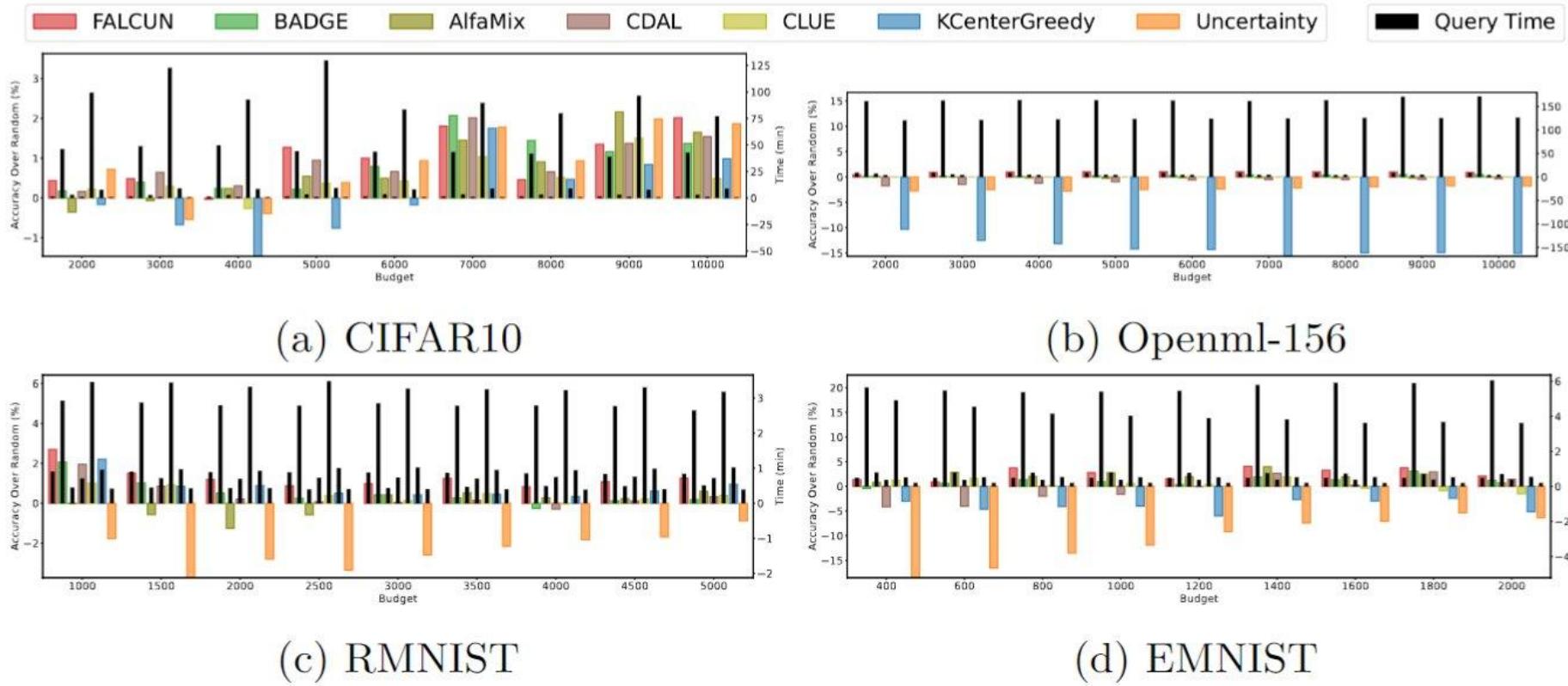
→ FALCUN selects instances (red circles) from classes the model cannot predict so well (dark areas) and less from classes with high accuracy (light areas)



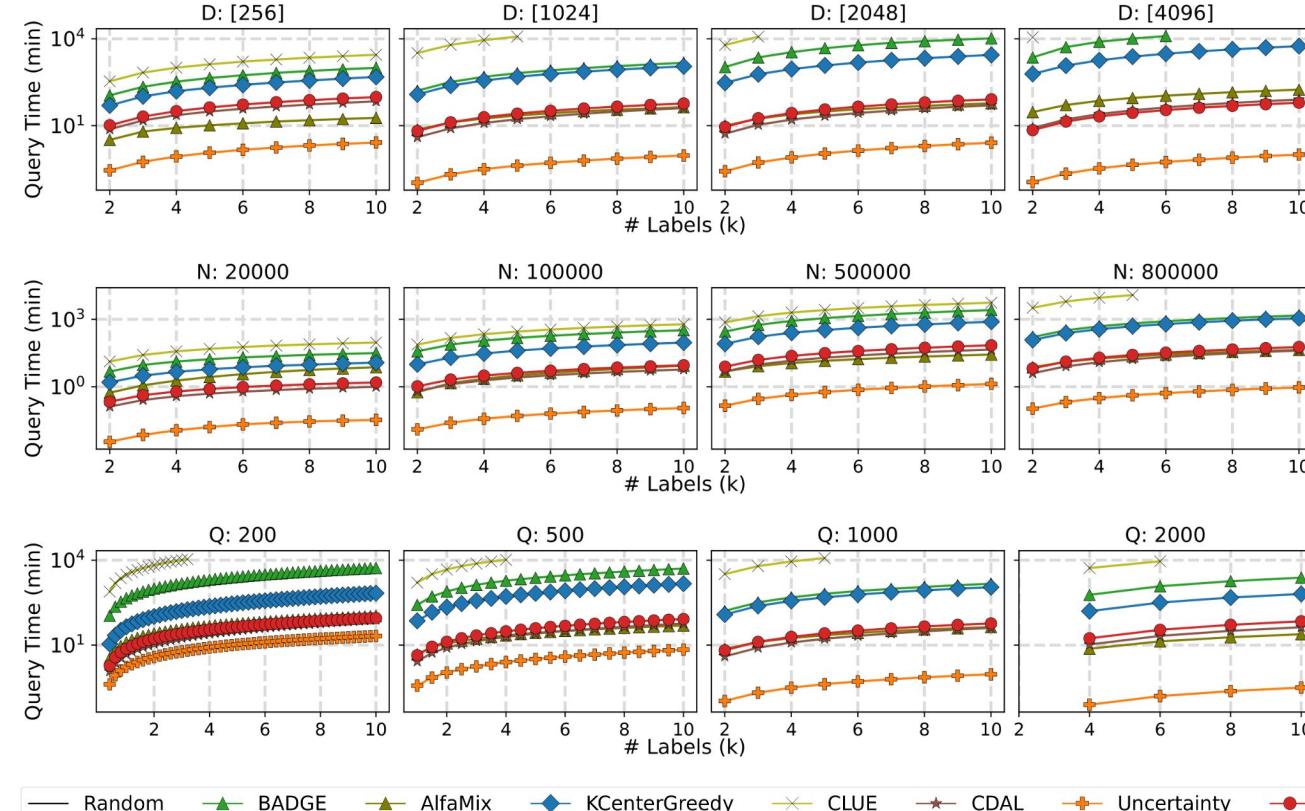
Results in short



Results - Label Efficiency



Results - Time Efficiency

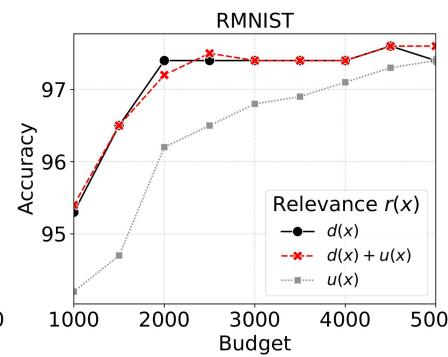
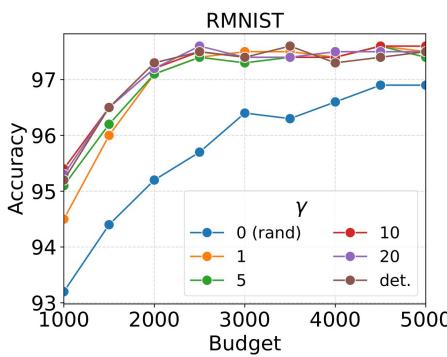
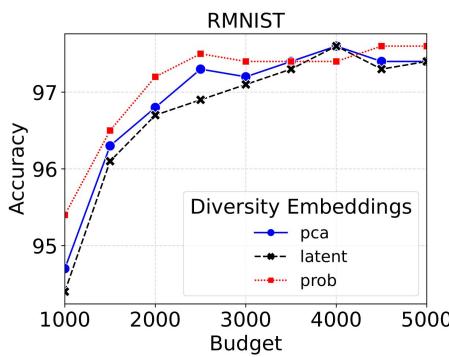
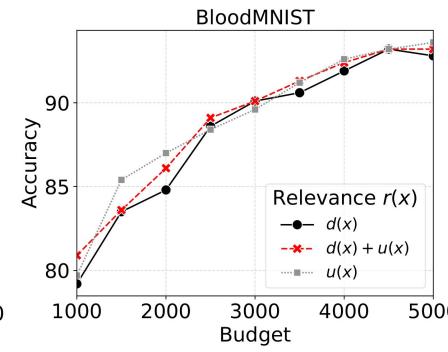
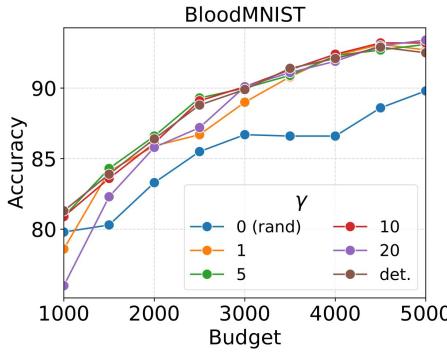
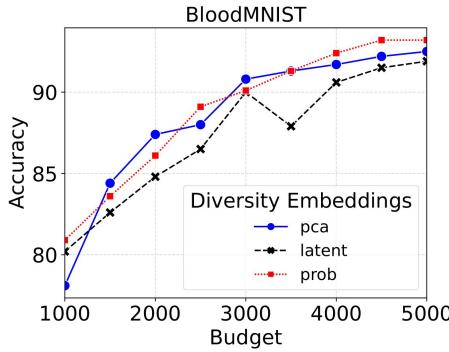


Fastest: Entropy

CLUE + BADGE + KCenterGreedy do not scale well to large dimensionality, large unlabeled pool, large query size

CDAL+ FALCUN+ Alfamix have similar order of magnitude, but Alfamix is more sensitive to dimensionality

Ablation



Diversity on probability space has strongest performance

$\Gamma \geq 5$ yields robust results

Diversity important on redundant instances

Overview

