Angular Gap: Reducing the Uncertainty of Image Difficulty through Model Calibration

Bohua Peng[1]

Mobarakol Islam [1]

Mei Tu [2]

1. Imperial College London   2. Samsung Research

Imperial College London

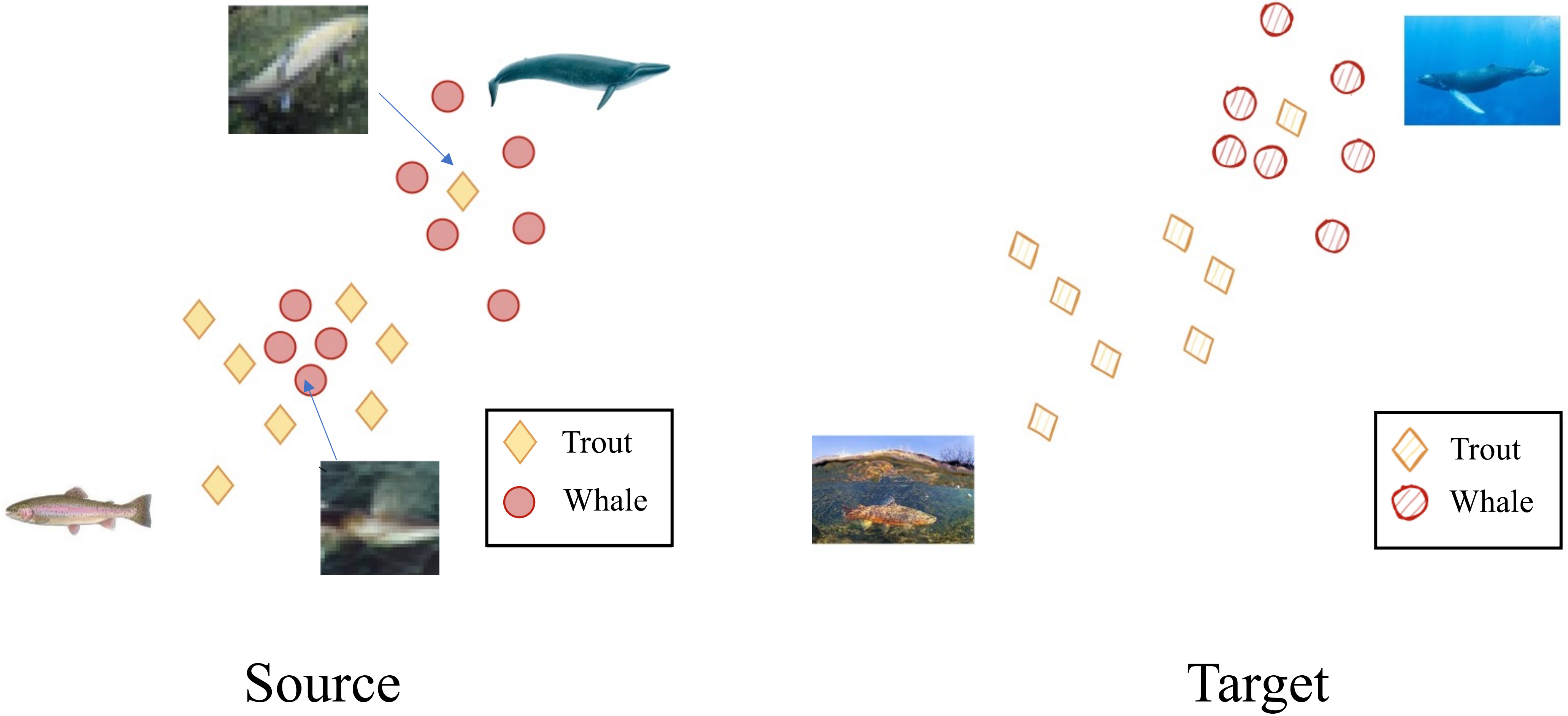
Samsung Research

Outline

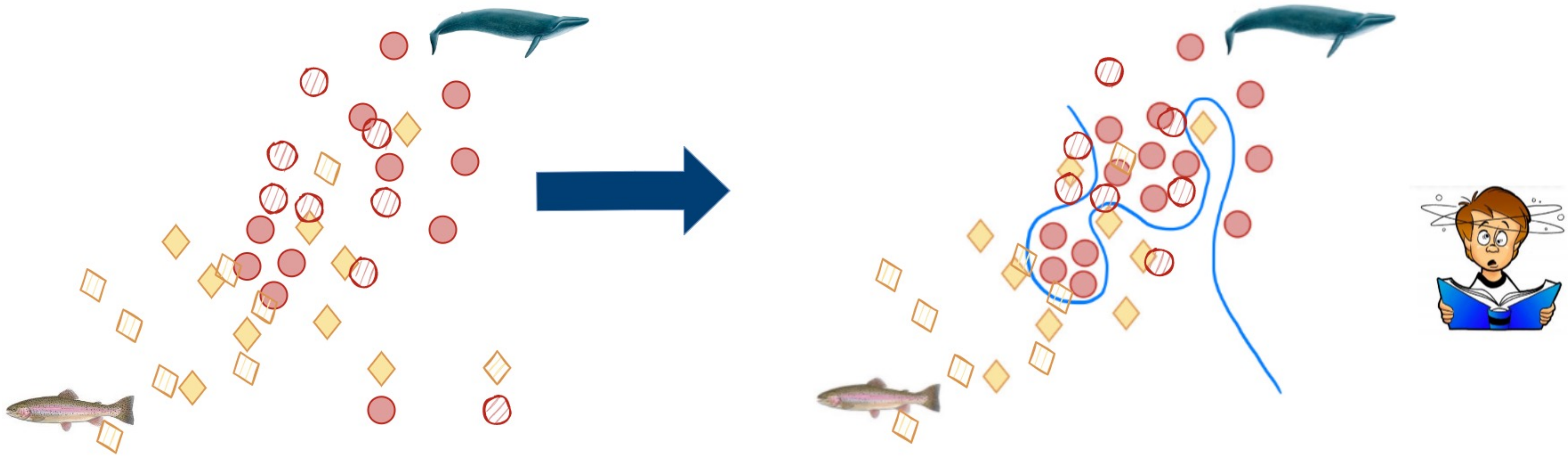
- Motivation
- Angular Gap framework
- Image difficulty estimator
- Curricula for domain adaptation
- Experiments
- Conclusion

Motivation

Capturing **image difficulty** & designing **curricula** for unsupervised **domain adaptation**



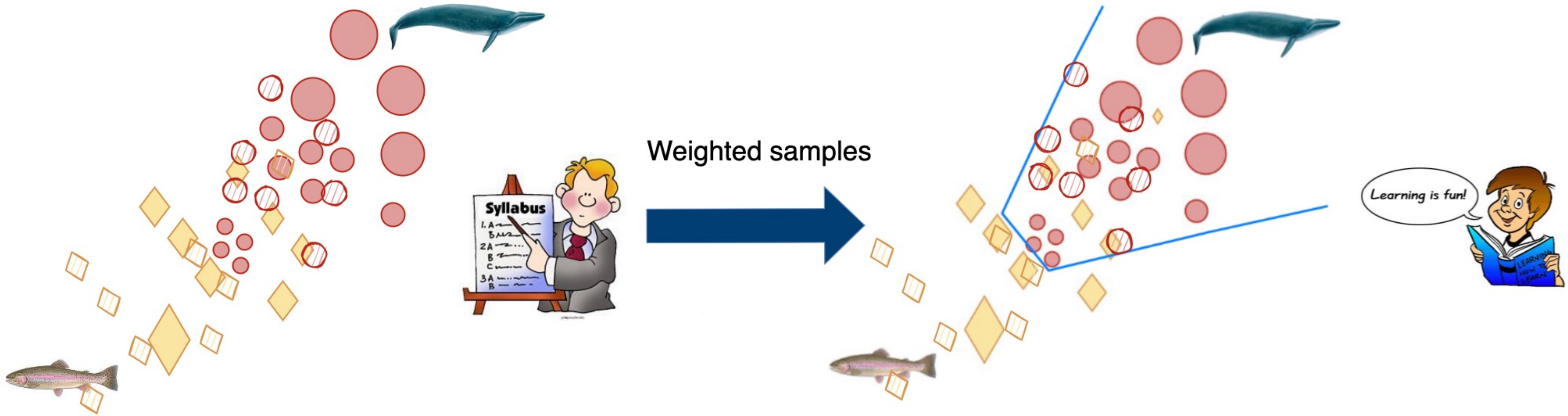
Motivation



Memorization

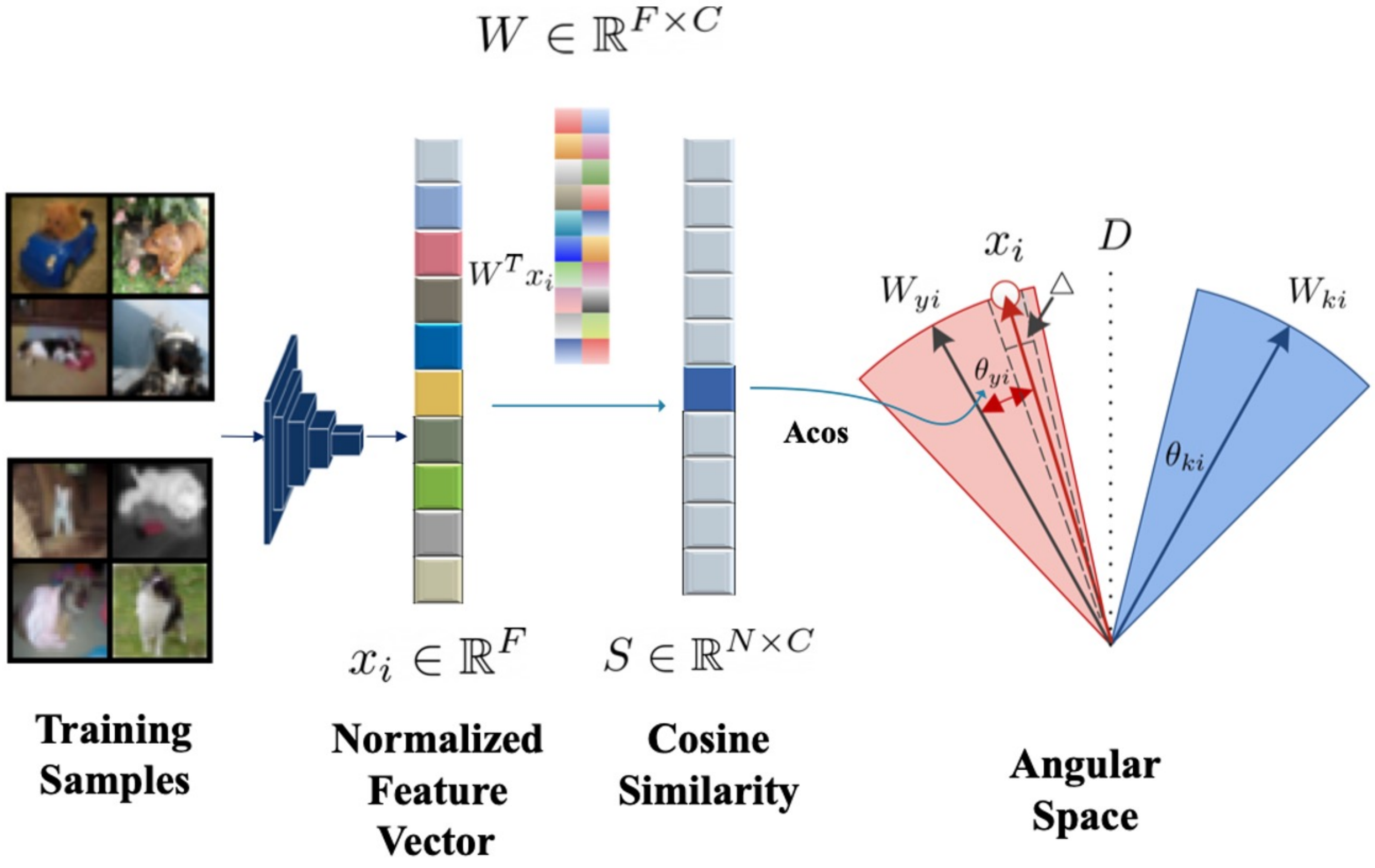
Motivation

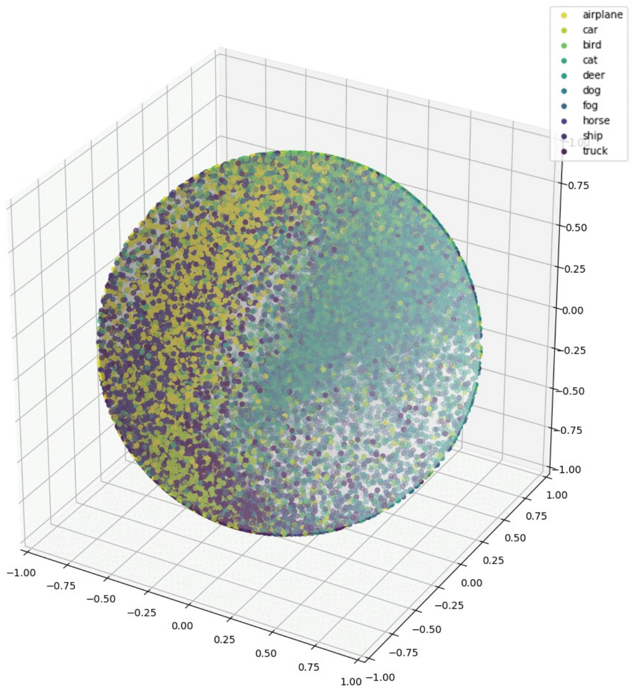
Learning with a **curriculum** to improve model **generalization**



How to represent image difficulty?

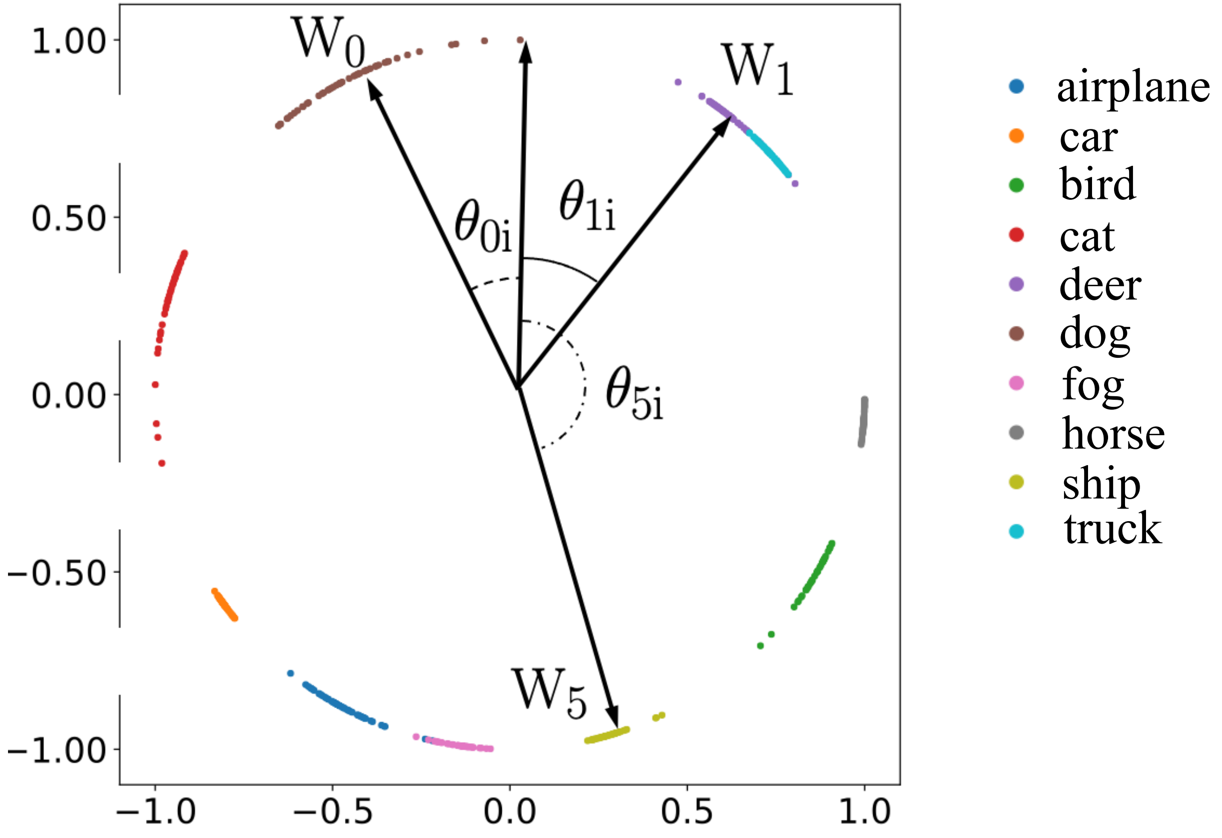
- Hyperspherical Learning (Liu, et al 2017), Angular Visual Hardness (Chen, 2020)

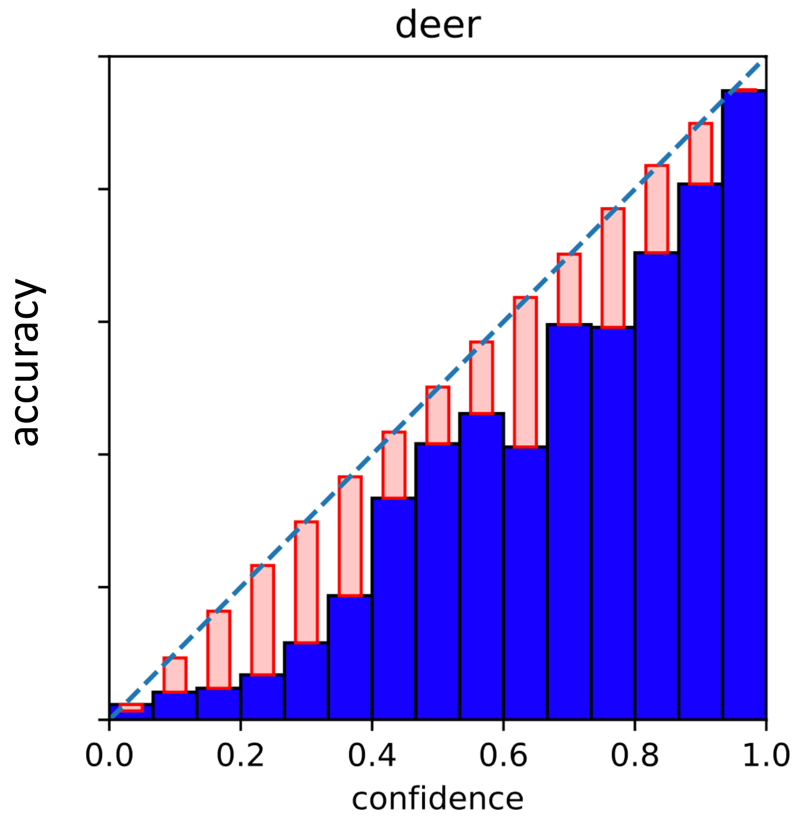




How to reduce uncertainty in difficulty estimation?

- Uncertain difficulty scores will ruin curriculum learning.

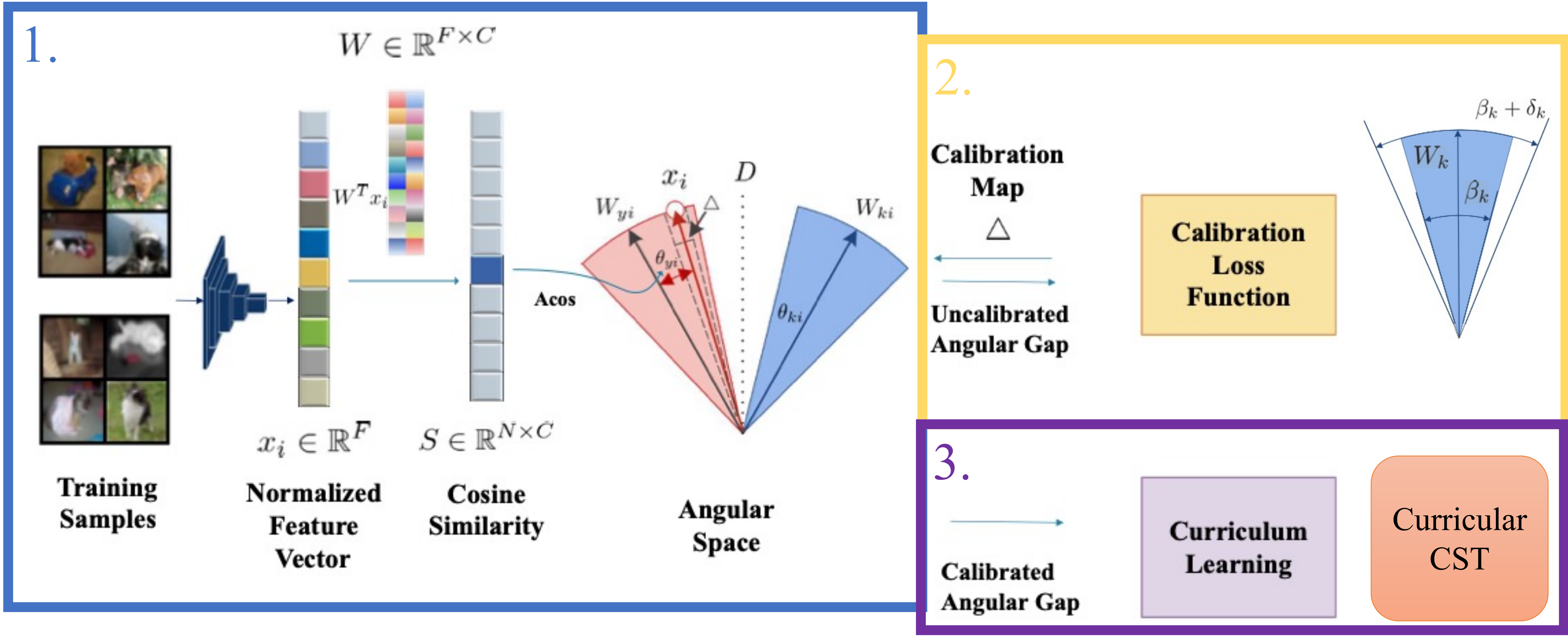




Main research questions:

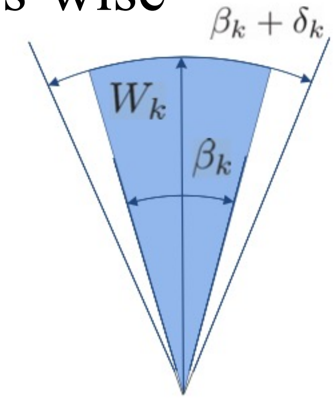
1. How to properly define image difficulty?
2. Will calibration techniques reduce uncertainty and improve downstream tasks with curriculum learning?
3. Can we automatically find a curriculum for domain adaptation tasks?

Angular Gap based curriculum learning

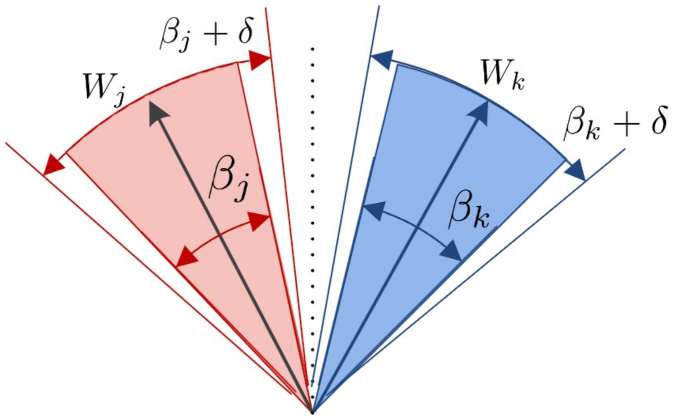


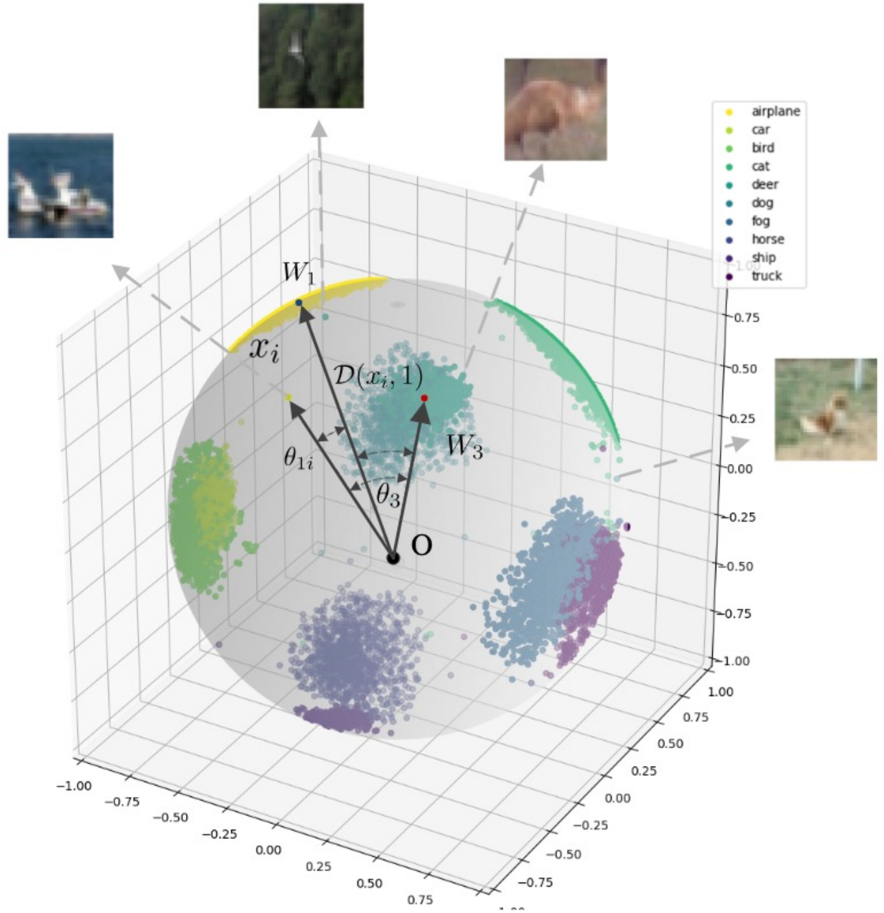
2.Multilevel model calibration

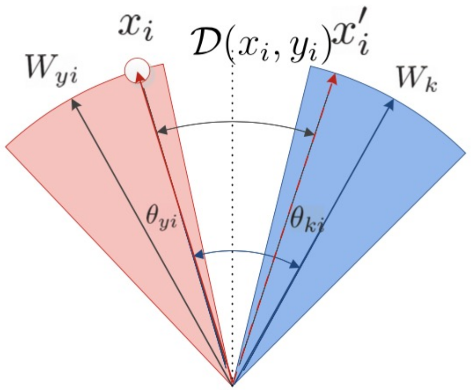
class-wise



global





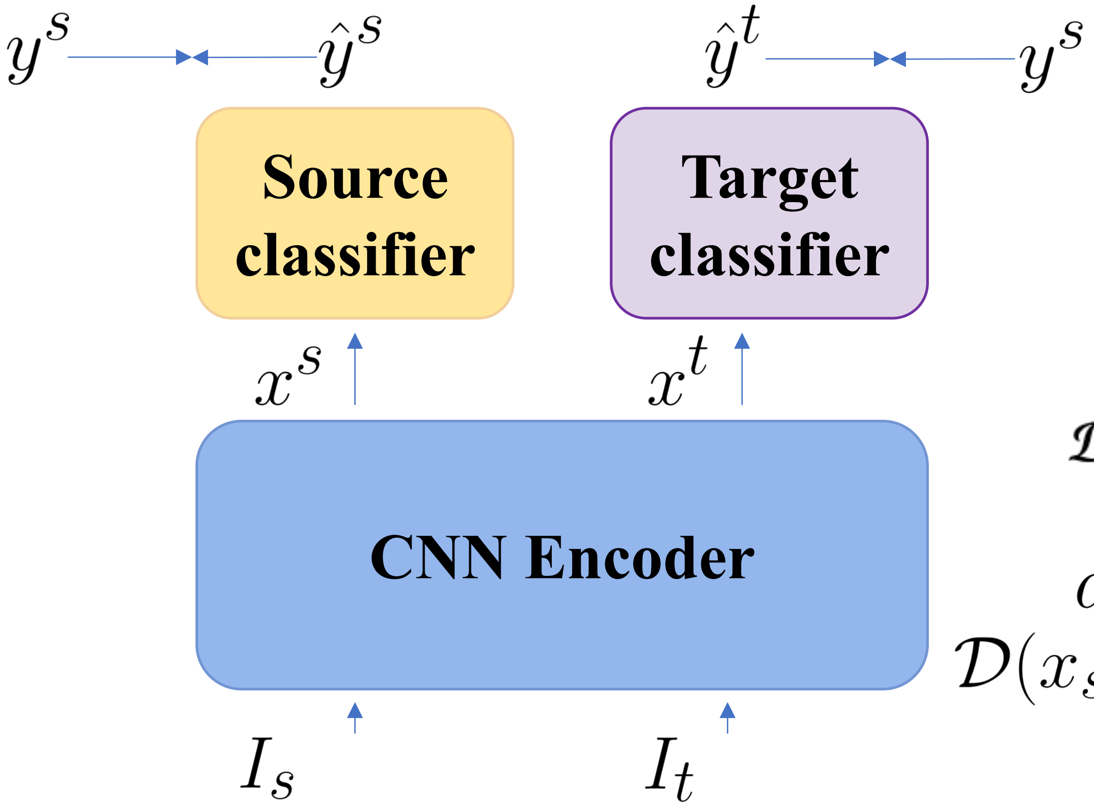


$$\mathcal{D}(\mathbf{x}, \mathbf{y}) = sim(\mathbf{x}, \mathbf{w}_y) - \arg\max_{k \neq y} sim(\mathbf{x}, \mathbf{w}_k)$$

$$sim(\mathbf{x}, \mathbf{w}_k) = \cos\theta_k = \frac{\mathbf{x}_i^T \mathbf{w}_k}{\|\mathbf{x}\| \, \|\mathbf{w}_k\|},$$

3.Find the curriculum d_s for distributional shift

Reducing cross-domain discrepancy with a curriculum for CST (Hong Liu, etc 2021)



$$L_{rev} \triangleq \sum_{s \in S} \sum_{t \in T} d_s \left\| sim(x_s, x_t)\hat{y}_t - y_s \right\|^2$$

$$d_s = \sigma(\lambda \cdot \mathcal{D}(x_s, y_s)),$$

$$\mathcal{D}(\mathbf{x}_s, \mathbf{y}_s) = sim(\mathbf{x}_s, \mathbf{w}_y) - \arg\max_{k \neq y} sim(\mathbf{x}_s, \mathbf{w}_k)$$

d_s is the curriculum found by grid search.
$\mathcal{D}(x_s, y_s)$ is the class-wise calibrated Angular Gap of the source domain samples.

Experiments

- HSF datasets (**limited resource**)
  - CIFAR10-H (Battleday, etc 2020)
  - ImageNetV2 (BenjaminRecht, etc 2019)

- Domain adaptation datasets
  - Office31 (Saenko, etc 2010)
  - VisDA2017 (Peng, etc 2017)

➢ Intrinsic Evaluation
  - Spearman rank correlation analysis (Charles Spearman)
  - Kendall Tau correlation analysis (Maurice Kendall 1938)
  - Expected Calibration Error (ECE)

➢ Extrinsic Evaluation
  - classification accuracy on CIFAR10
  - Performance on domain adaptation

Evaluation

Table1. Correlation analysis of example difficulties and HSF on CIFAR10-H. Performance of corresponding curriculum learning.

Methods	Spearman's rank	Kendall's Tau	ECE	Top-5 acc.	Top-1 acc.
Maximum Confidence	0.266±0.006	0.148±0.004	11.3±0.2	94.5±0.3	74.8±2.1
Maximum Confidence$_{\text{TS}}$	0.273±0.004	0.145±0.004	9.1±0.2	94.5±0.3	75.3±2.1
Classfication Margin[39]	0.279±0.006	0.142±0.004	11.3±0.2	94.6±0.3	75.3±2.0
Classfication Margin$_{\text{TS}}$	0.283±0.006	0.242±0.004	9.0±0.2	94.9±0.3	75.7±1.3
MC-Dropout[10]	0.256±0.007	0.176±0.006	9.4±0.4	95.7±0.3	77.0±0.5
AVH[5]	0.368±0.006	0.258±0.004	8.2±0.2	98.2±0.3	81.2±1.0
AVH$_{\text{Global}}$	0.376±0.003	0.263±0.003	<u>7.5±0.2</u>	98.6±0.2	81.3±0.7
AVH$_{\text{Class-wise}}$	0.377±0.003	0.265±0.002	**7.4±0.2**	98.6±0.2	81.4±0.6
Forgetting Events[39]	0.260±0.003	0.187±0.002	11.5±0.5	98.0±0.4	78.9±1.2
C-score[18]	0.316±0.001	0.243±0.001	9.8±0.3	**99.0±0.1**	**82.4±0.4**
Prediction Depth[1]	0.290±0.001	0.183±0.001	9.8±0.3	98.5±0.2	81.2±0.4
Angular Gap (Ours)	0.378±0.003	0.265±0.003	8.2±0.2	98.6±0.2	82.0±0.6
Angular Gap$_{\text{Global}}$	<u>0.382±0.003</u>	<u>0.268±0.002</u>	<u>7.5±0.2</u>	98.8±0.2	82.3±0.4
Angular Gap$_{\text{Class-wise}}$	**0.384±0.002**	**0.269±0.002**	**7.4±0.2**	<u>98.9±0.2</u>	**82.4±0.4**

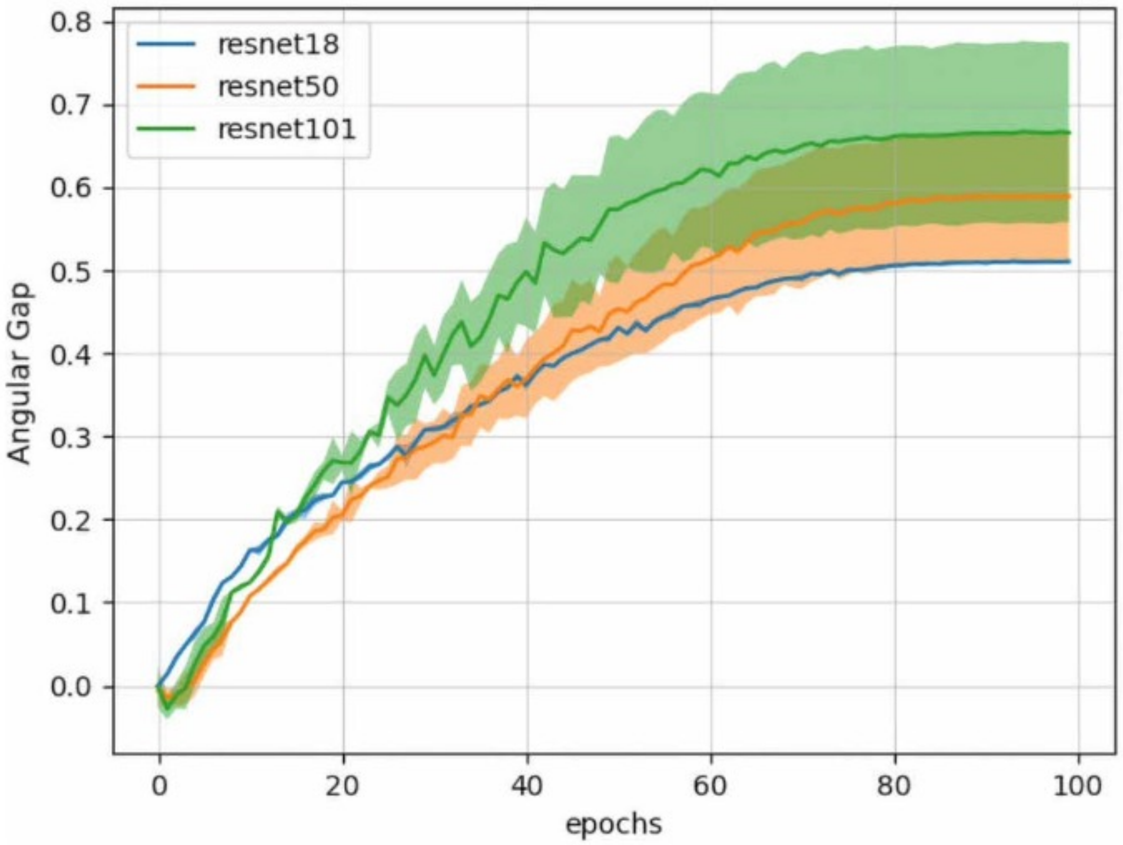
Evaluation

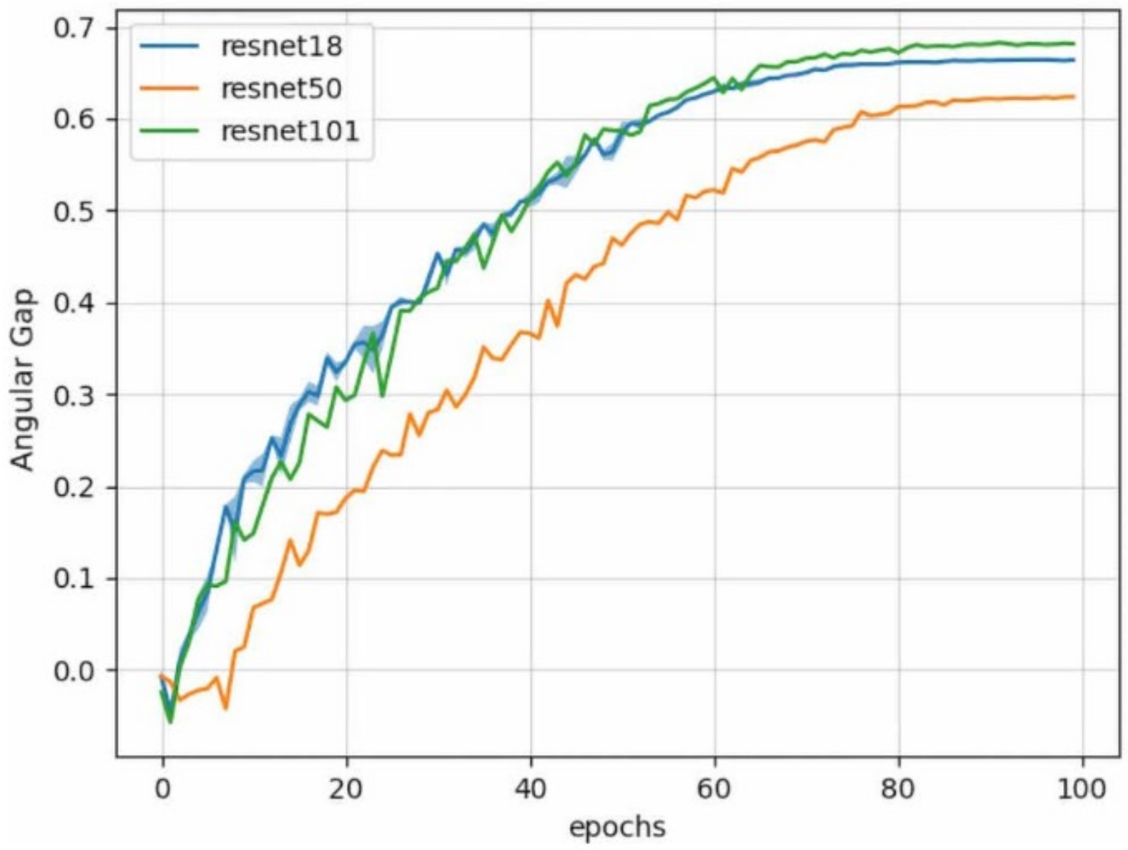
Table 2: Accuracy (%) for sythetic-to-real on VisDA2017

Method	Acc.	Method	Acc.
DANN[11]	55.3	CBST[46]	76.4
DAN[27]	61.1	CRST[47]	78.1
MSTN[43]	65.0	FixMatch[38]	76.7
JAN[28]	65.7	CST[25]	79.9
DSAN[45]	74.8	FixBi[29]	<u>87.2</u>
Curricular DSAN	75.4	Curricular CST	**88.1**

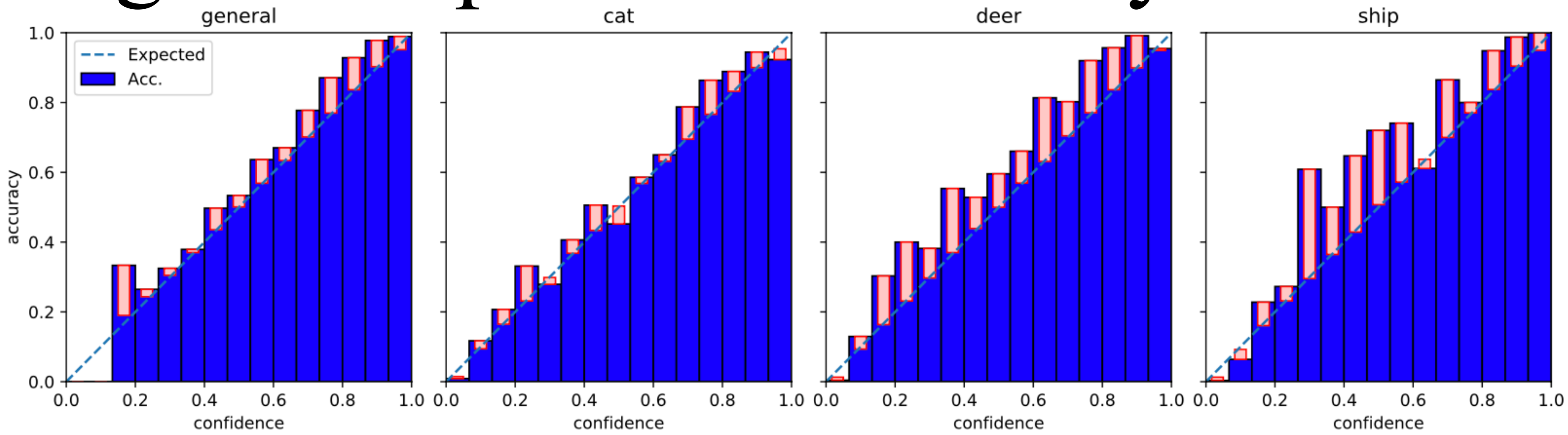
Why does calibrated Angular Gap work?

Model calibration reduces the uncertainty and improves robustness of image difficulty estimation.

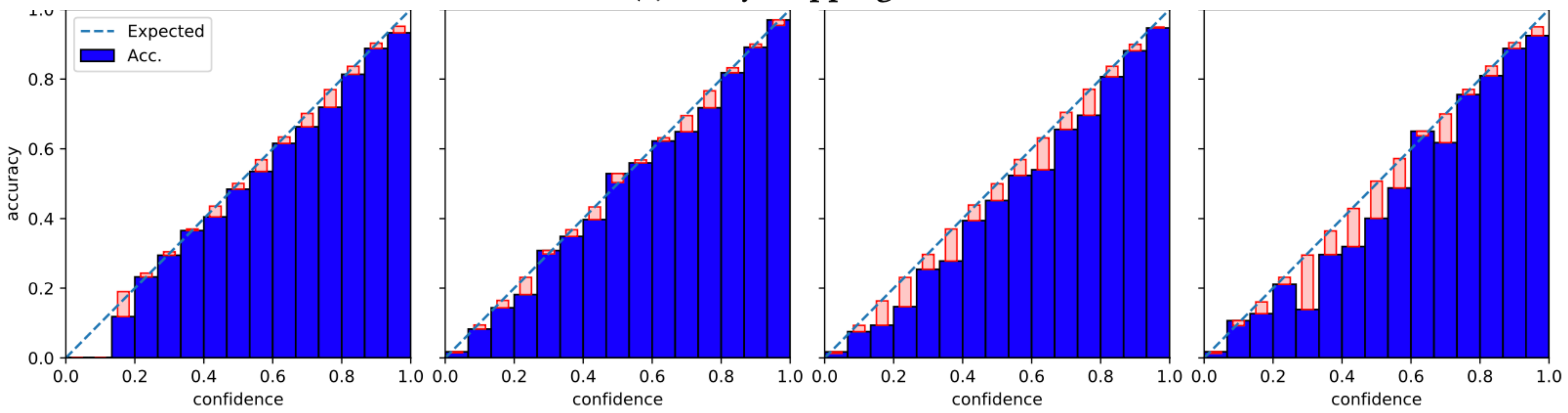




Angular Gap uses trustworthy similarities



(a) Early stopping



(b) Class-wise calibration with FN

Conclusion

- We propose Angular Gap to estimate image difficulty and reduce its uncertainty with class-wise calibration.
- Our results benefit from feature normalization of hyperspherical learning.
- We build an Angular Gap based curriculum with CST for UDA.
- Model calibration benefits image difficulty estimation. Simply scaling up the model leads to unstable difficulty estimation.

Thank you