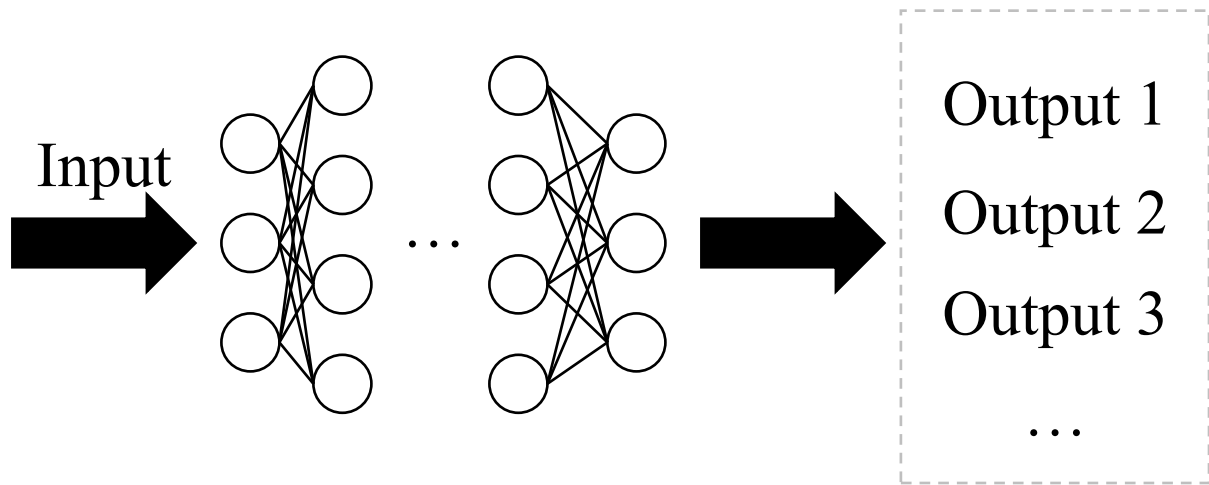


Uncertainty Estimation via Hyperspherical Confidence Mapping

Motivation

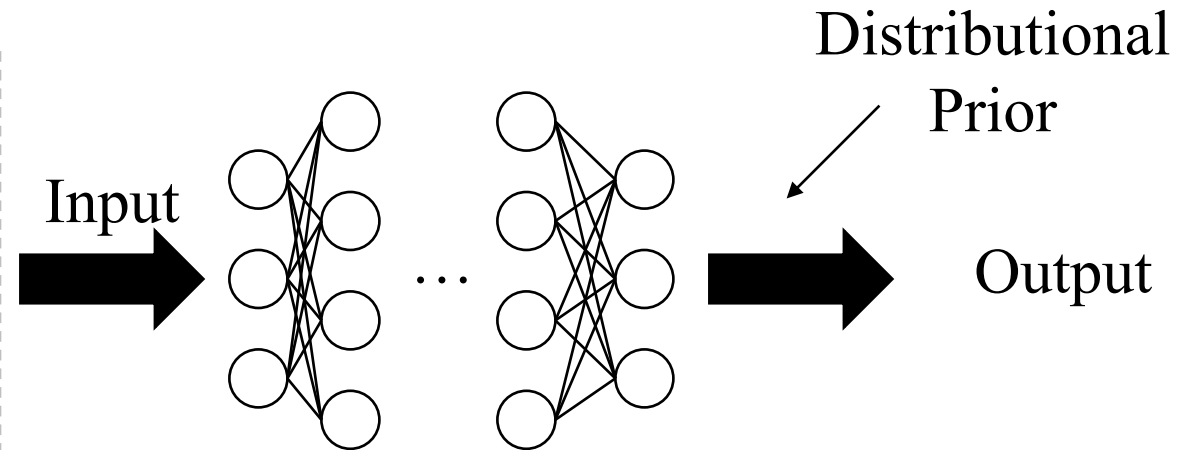
Why Uncertainty Estimation Matters

Sampling based



- Multiple Forward Passes
- Computationally Expensive

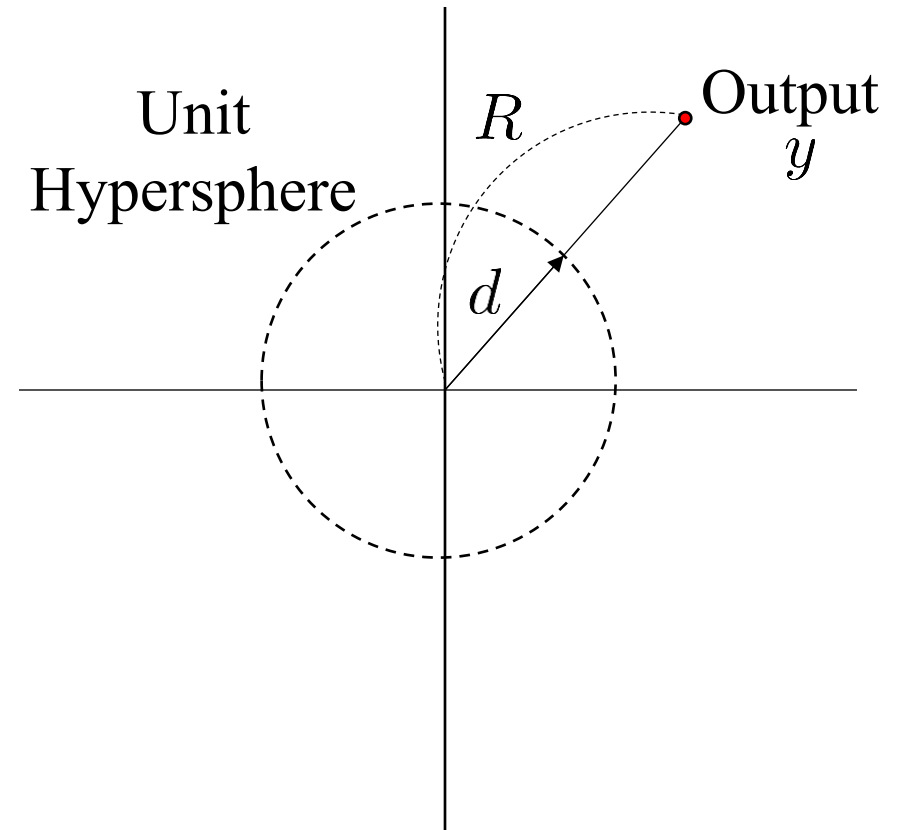
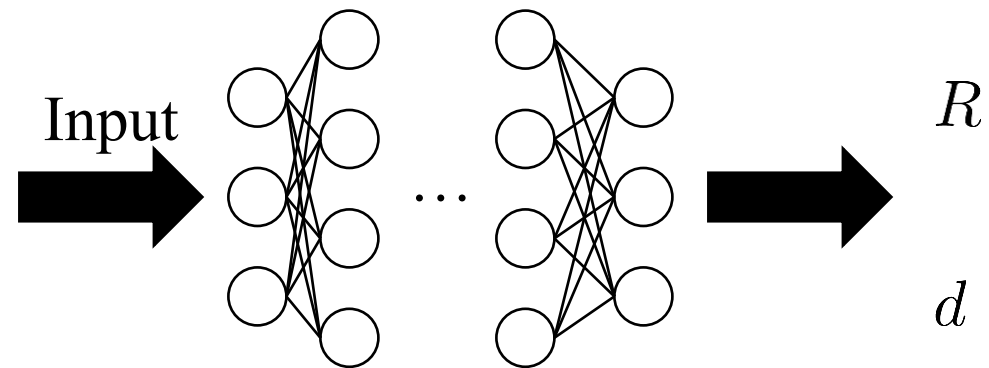
Distribution based



- Assume specific form
- May not capture multimodal uncertainty

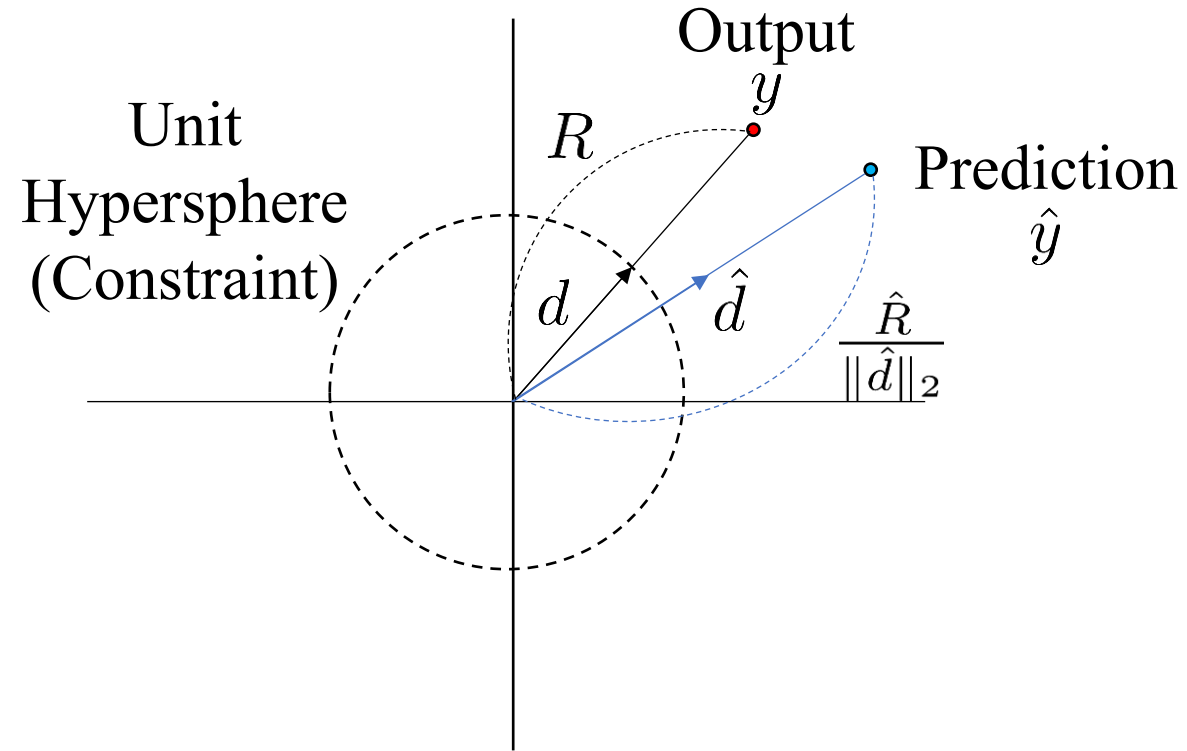
Key Idea

Hyperspherical Confidence Mapping (HCM)



Core Insight

Uncertainty as Constraint Violation



$$\text{Uncertainty Score } u := \hat{R} \left| \|\hat{d}\|_2 - 1 \right|$$

Theory

Lower Bound on Prediction Error

$$\hat{R}\|d - \hat{d}\|_2 - |R - \hat{R}| \leq \|y - \hat{y}\|_2 \leq \hat{R}\|d - \hat{d}\|_2 + |R - \hat{R}|$$

Gaussian Noise Estimator

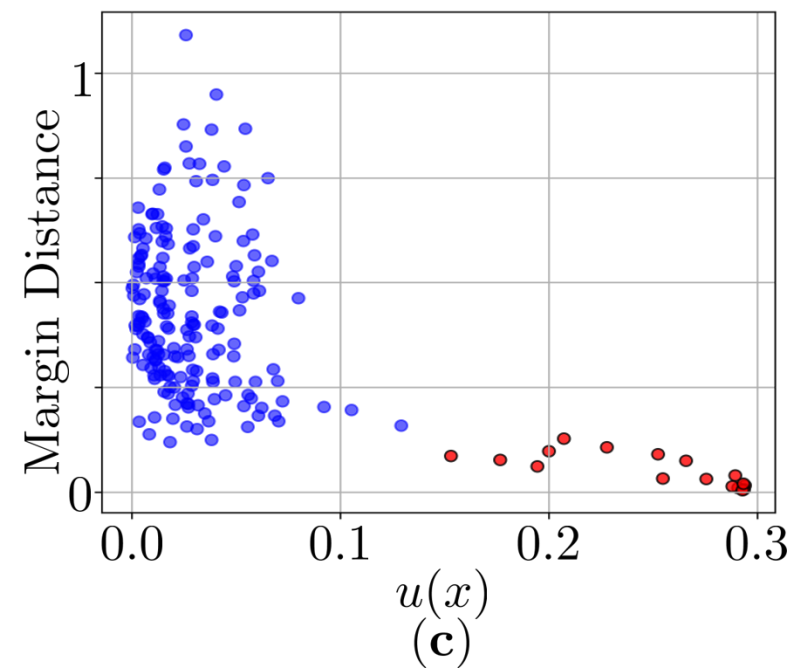
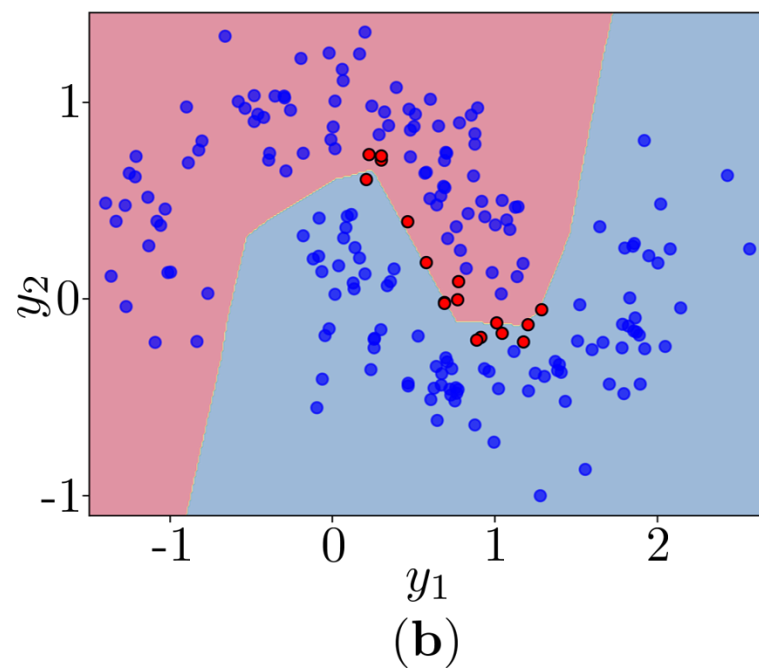
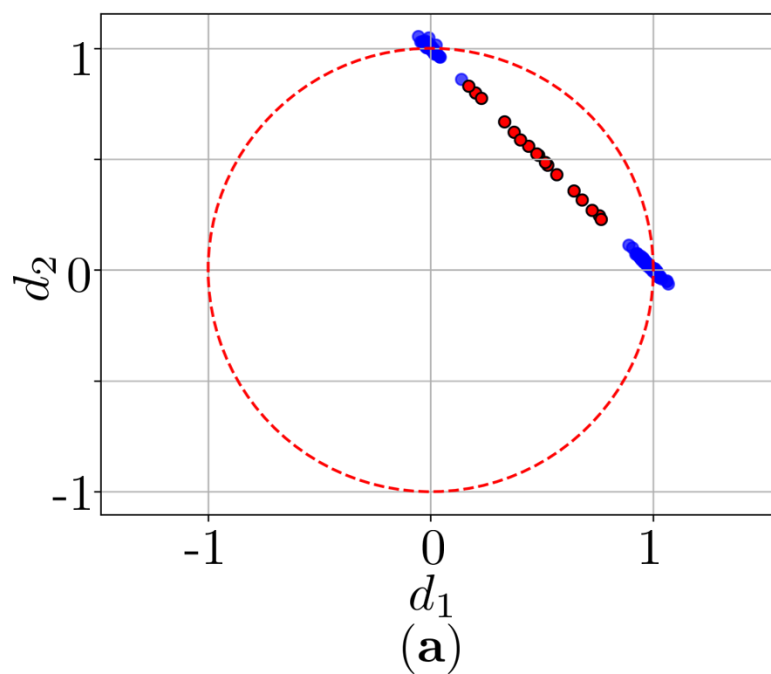
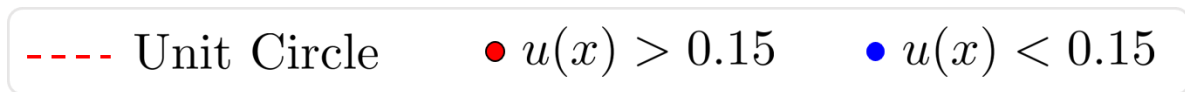
If $y = g(x) + \xi$, ($\xi \sim \mathcal{N}(0, \sigma^2)$), $y \in \mathbb{R}^D$

$$\hat{\sigma}^2(x) := \frac{1}{D-1}u(x) \left(\hat{R}(x)(1 + \|\hat{d}(x)\|_2) \right)$$

$$\hat{\sigma}^2(x) = \sigma^2 + O\left(\frac{(D+2)(D+4)\sigma^4}{(D-1)\|g(x)\|_2^2} \right)$$

Experimental Results

Two Moons



Experimental Results

OOD Detection

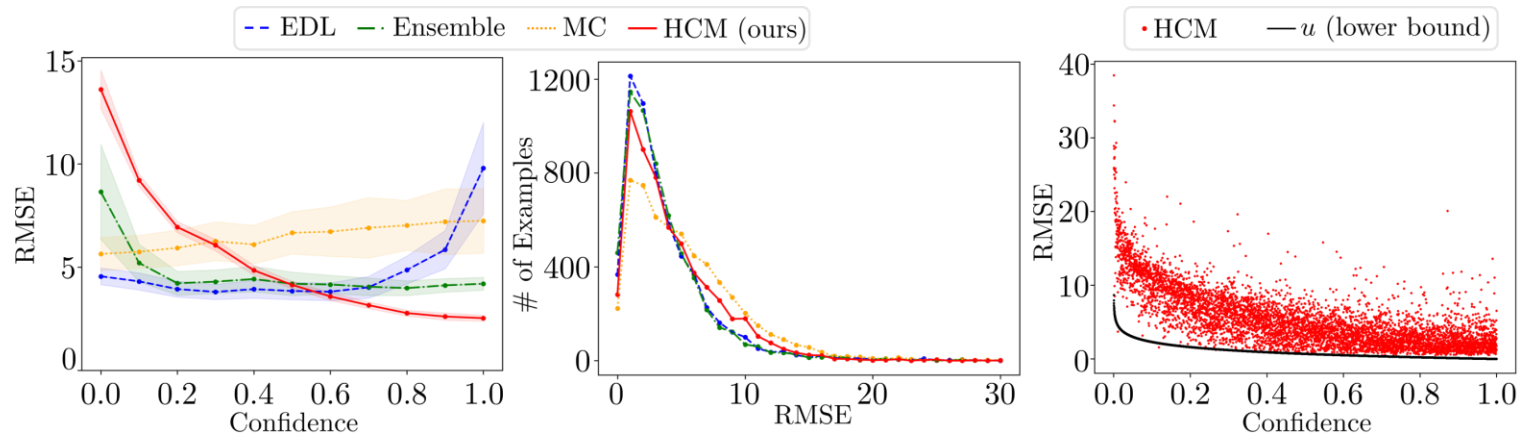
OOD Type	MSP	Ensembles	MC	ODIN	Energy	MDS	KNN	ViM	fDBD	NCI	HCM	HCM mix
Near OOD	86.73	88.89	85.21	85.49	87.52	84.91	88.07	84.20	88.87	86.49	82.23	87.90
Far OOD	88.96	90.86	90.50	88.46	89.36	91.90	92.59	90.52	69.52	92.49	86.45	90.12
AVG	87.85	90.15	88.33	87.23	88.62	89.34	90.97	88.12	76.96	90.36	85.04	89.44

➤ *Efficiency benchmark of OOD detection methods*

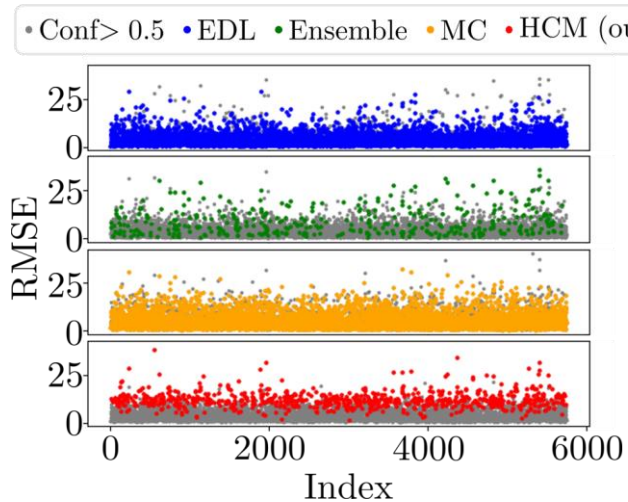
Method	Extra Cost	Latency (ms/img)	Throughput (img/s)	Peak Mem (MB)
MSP	Softmax only	0.15–0.20	6k–7k	227
Ensemble (M=5)	5× forward passes	0.25–0.38	2.6k–4.3k	577
MC Dropout	50× stochastic passes	1.6–1.8	~600	227
ODIN	MSP + backward wrt input	0.36–0.66	1.5k–2.8k	452
Energy	log-sum-exp energy	0.16–0.33	3k–6k	227
MDS	feature + Mahalanobis score	0.20–0.36	4.9k–5.9k	228
KNN	feature + KNN search	0.51–0.66	1.5k–2.0k	227
VIM	feature + covariance eigenspace	0.18–0.20	5.4k–5.5k	227
fDBD	feature + distance-based score	0.21	4.7k–6.3k	228
NCI	feature + canonical corr score	0.37	2.7k–6.5k	228
HCM (ours)	magnitude + direction calc	0.19	5k–6.8k	226

Experimental Results

Industrial High-Dimension Regression



➤ *Samples with Confidence below 0.1*



- Baselines show weak correlation with error
- HCM captures high-error samples

Thank you for watching