Topology-matching Normalizing Flows for Out-of-Distribution Detection in Robot Learning

Jianxiang Feng[1,4], Jongseok Lee[2,3], Simon Geisler[1], Stephan Günnemann[1], Rudolph Triebel[2,3]

[1]Department of Informatics, Technical University of Munich (TUM)
[2]Institute of Robotics and Mechatronics, German Aerospace Center (DLR)
[3] Department of Informatics, Karlsruhe Institute of Technology (KIT) [4] Agile Robots AG

Motivation

Out-of-distribution (OOD) detection of a trustworthy open-world robot should be:

- ❖ **accurate;**
- ❖ **cost-efficient;**
- ❖ **easy-to-use.**

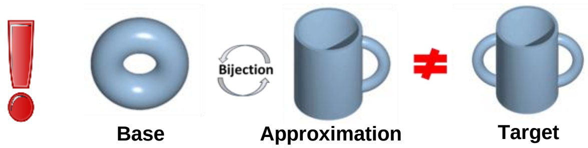
⇒ **Normalizing flows (NFs)**.

Aerial manipulators in the wild

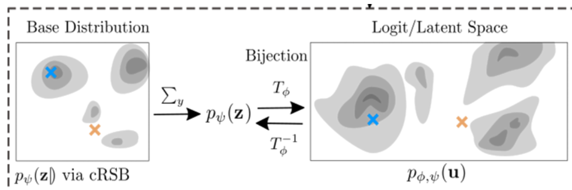
Assistive robots in the household

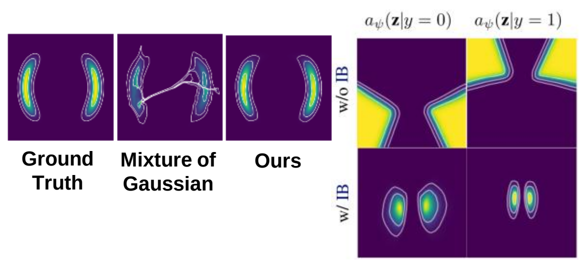
Problem of NFs:

- **hard** to model distributions with **complex topology**.



Topology-matching Normalizing Flows





★ **Conditional Resampled Base Distribution (cRSB)**

$$p_\psi(\mathbf{z}|y) = (1-\alpha_T)\frac{a_\psi(\mathbf{z}|y)\pi(\mathbf{z})}{Z_y} + \alpha_T\pi(\mathbf{z}),$$

where $a_\psi : \mathcal{R}^d \to [0,1]^C$ and $\alpha_T = (1-Z_y)^{T-1}$.

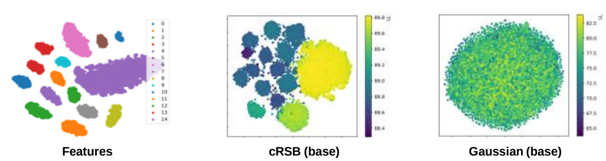
★ **Information Bottleneck (IB) Objective**

$$\mathcal{L}_{\text{IBNF}} = CI(U, Z_\epsilon) - \beta CI(Z_\epsilon, Y)$$

$$CI(U, Z_\epsilon) = \mathbb{E}_{p(\mathbf{u}),p(\epsilon)}\left[-\log\sum_{y'}(p_\psi(\mathbf{z}_\epsilon|y')) - \log|\det(J_{T_\phi^{-1}}(\mathbf{u}+\epsilon))|\right],$$

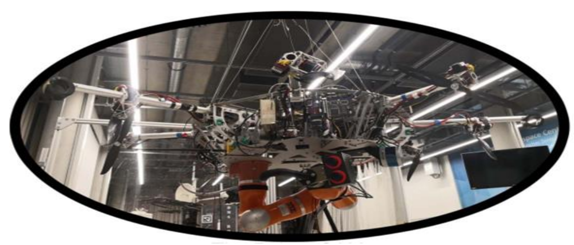
$$CI(Z_\epsilon, Y) = \mathbb{E}_{p(y)}\left[\log\frac{p_\psi(\mathbf{z}_\epsilon|y)p(y)}{\sum_{y'}(p_\psi(\mathbf{z}_\epsilon|y')p(y'))}\right].$$

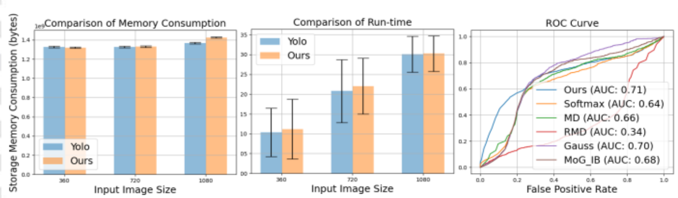
Benchmark Results



	Pascal-VOC-OS				MS-COCO-OS			
	AUROC	TPR at 5%FPR	TPR at 10%FPR	TPR at 20%FPR	AUROC	TPR at 5%FPR	TPR at 10%FPR	TPR at 20%FPR
Softmax	0.901	60.1	72.8	83.1	0.882	61.3	70.6	78.1
Entropy	0.905	59.8	72.9	82.9	0.903	61.2	70.6	80.2
MD [62]	0.9	54.1	68.8	83.3	0.902	57.2	71.4	85.5
RMD [63]	0.838	15.2	28.4	77.4	0.531	1.7	2.6	7.1
Ensemble Softmax [58]	0.885	47.8	72.6	83.1	0.898	66.2	73.5	82.3
Ensemble Entropy [58]	0.887	47.8	72.5	83.1	0.906	66.2	73.5	82.3
GMMDet [25]	0.931	70.7	*80.5*	*89.3*	0.924	69.5	80.2	87.9
Flows Gaussian	0.915 ± 0.002	72.2 ± 0.75	77.8 ± 0.89	86.1 ± 0.67	0.924 ± 0.001	68.2 ± 0.73	81.2 ± 0.61	89.4 ± 0.04
Flows MoG	0.919 ± 0.002	69.0 ± 2.4	77.0 ± 2.5	86.5 ± 1.2	0.925 ± 0.001	68.3 ± 0.30	80.5 ± 0.50	89.6 ± 0.05
Flows RSB [13]	0.924 ± 0.003	72.8 ± 0.88	79.3 ± 1.0	87.1 ± 0.82	0.925 ± 0.001	68.6 ± 0.87	81.3 ± 0.31	89.5 ± 0.34
Flows MoG_CLS [27]	0.923 ± 0.001	69.2 ± 1.5	78.2 ± 1.3	88.5 ± 0.82	*0.930* ± 0.001	68.5 ± 0.73	*82.2* ± 0.31	*89.7* ± 0.30
Flows MoG_IB [8]	*0.934* ± 0.002	73.1 ± 1.3	79.6 ± 0.6	87.8 ± 0.2	0.924 ± 0.002	*71.1* ± 0.9	79.6 ± 0.46	88.6 ± 0.63
Flows cRSB_CLS	0.919 ± 0.001	72.5 ± 0.37	78.8 ± 0.27	86.8 ± 0.42	0.924 ± 0.001	68.3 ± 0.14	81.1 ± 0.30	89.3 ± 0.18
Flows cRSB_IB (ours)	**0.946** ± 0.003	**78.5** ± 0.97	**84.0** ± 0.83	**90.8** ± 0.76	**0.934** ± 0.002	**73.3** ± 2.0	**84.3** ± 0.40	**91.3** ± 0.28

Real Robot Deployment





Takeaway

- ❖ We propose to mitigate the fundamental topological mismatch problem in Normalizing Flows for effective OOD detection.
- ❖ We achieve this with the expressive cRSB and the IB objective. Extensive experiments demonstrate superior performance both quantitatively and qualitatively.
- ❖ The resultant NF-based OOD detector is lightweight and compatible with numerous existing object detectors.