Geometry-Aware Post-Hoc Uncertainty Quantification in Operator Learning

Oriol Vendrell-Gallart
PhD Student
University of California, Irvine
ovendrel@uci.edu

Nima Negarandeh
PhD Student
University of California, Irvine
nnegaran@uci.edu

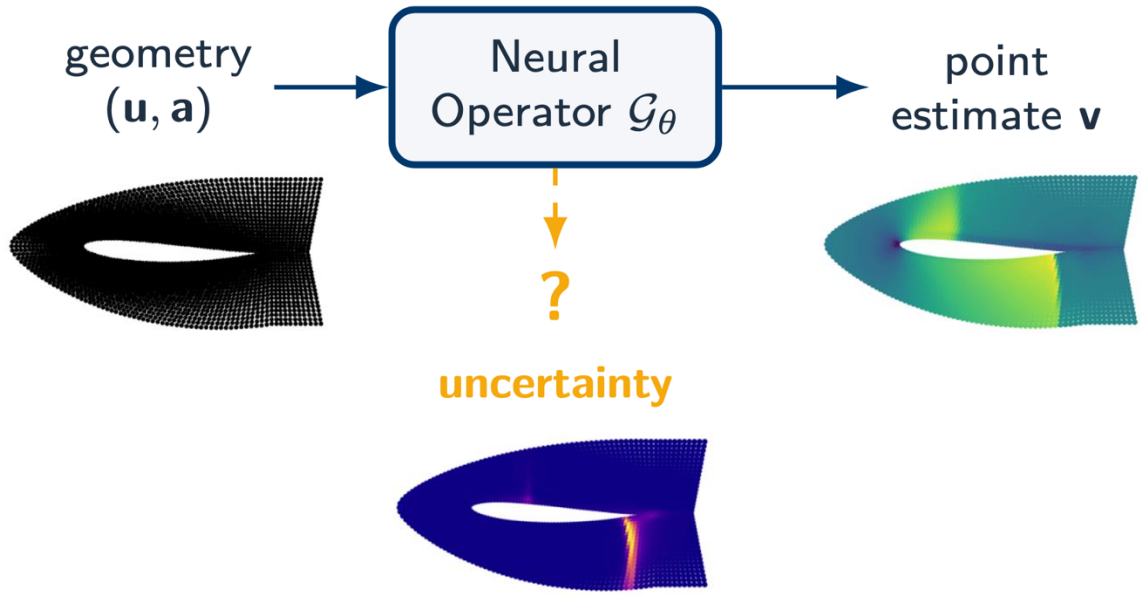
Ramin Bostanabad
Associate Professor
University of California, Irvine
raminb@uci.edu

Sydney, December 2026

NEURAL OPERATORS AS PDE SURROGATES

- Neural operators replace expensive solvers across climate, fluids, plasma and solid mechanics.
- Modern **geometry-aware** backbones (e.g. Transolver) run natively on **point clouds**.
- But they are trained **deterministically**:

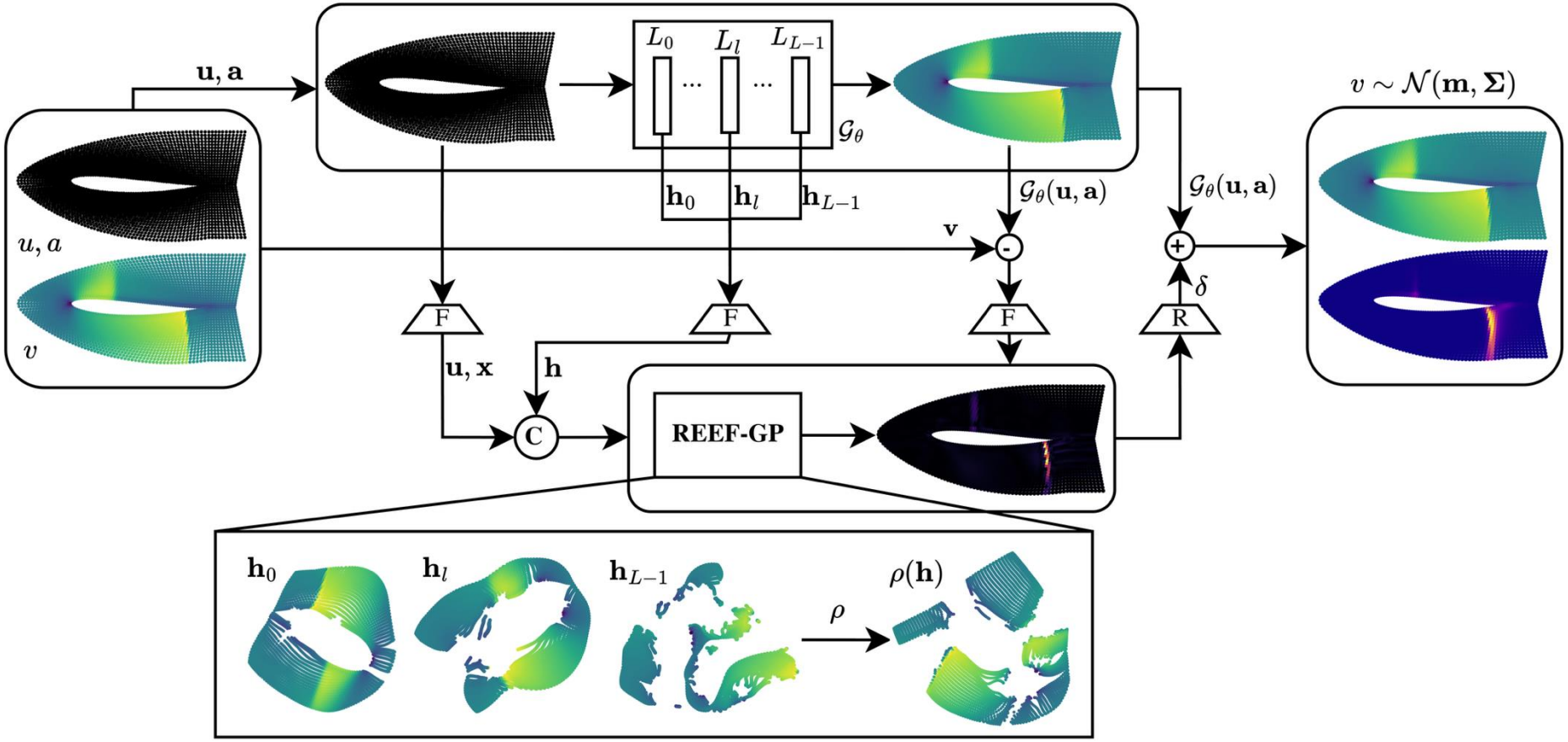
  one input → one output



Why this matters?

Scientific use and downstream tasks (active data collection, design, safety) need to know ***when to trust*** the surrogate, especially under ***geometric variability***.

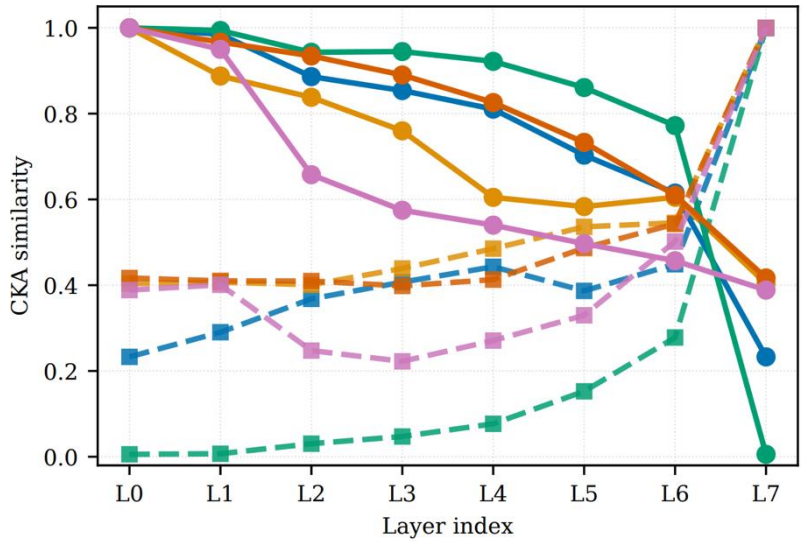
REEF-GP AT A GLANCE

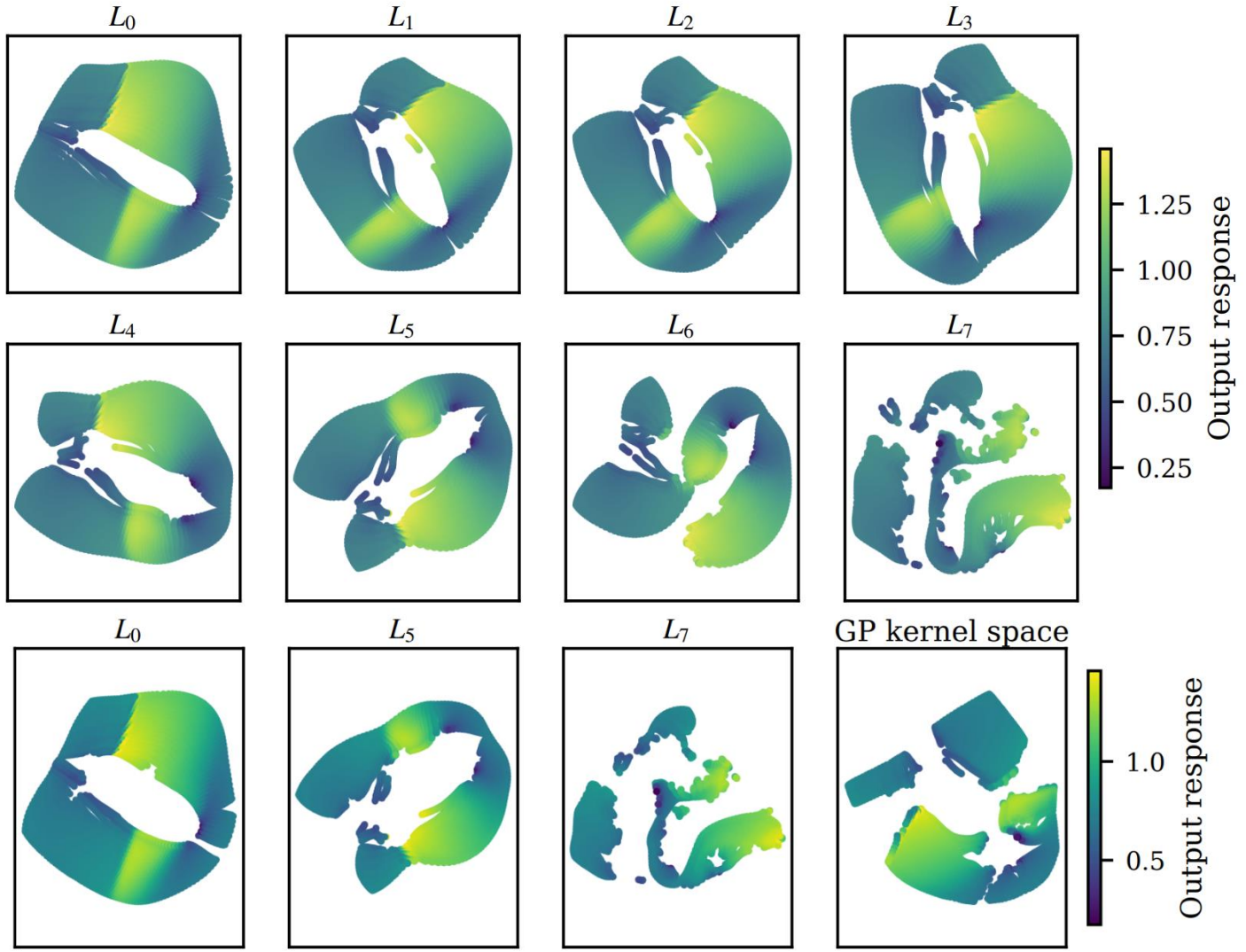


Frozen $\mathcal{G}_\theta$ maps $(\mathbf{u}, \mathbf{a}) \rightarrow \mathbf{v}$. REEF-GP models its **residual**: **transforms** (ρ) the operator's geometry-aware embeddings ($\mathbf{h_0}, \mathbf{h_l}, \mathbf{h_{L-1}}$) into a kernel space, and adds the discrepancy $\boldsymbol{\delta}$ to yield an output distribution. Note how the kernel space **tears exactly at the shock fronts**.

WHY INTERNAL LAYERS?

- Transolver's blocks act like a **discrete flow** (connection to Kernel Flows).
- Intermediate layers carry **richer** information than the final readout. We pick **3 CKA-distinct layers**.





Houman Owhadi and Gene Ryan Yoo. Kernel flows: From learning kernels from data into the abyss. Journal of Computational Physics, 389:22–47, 2019.

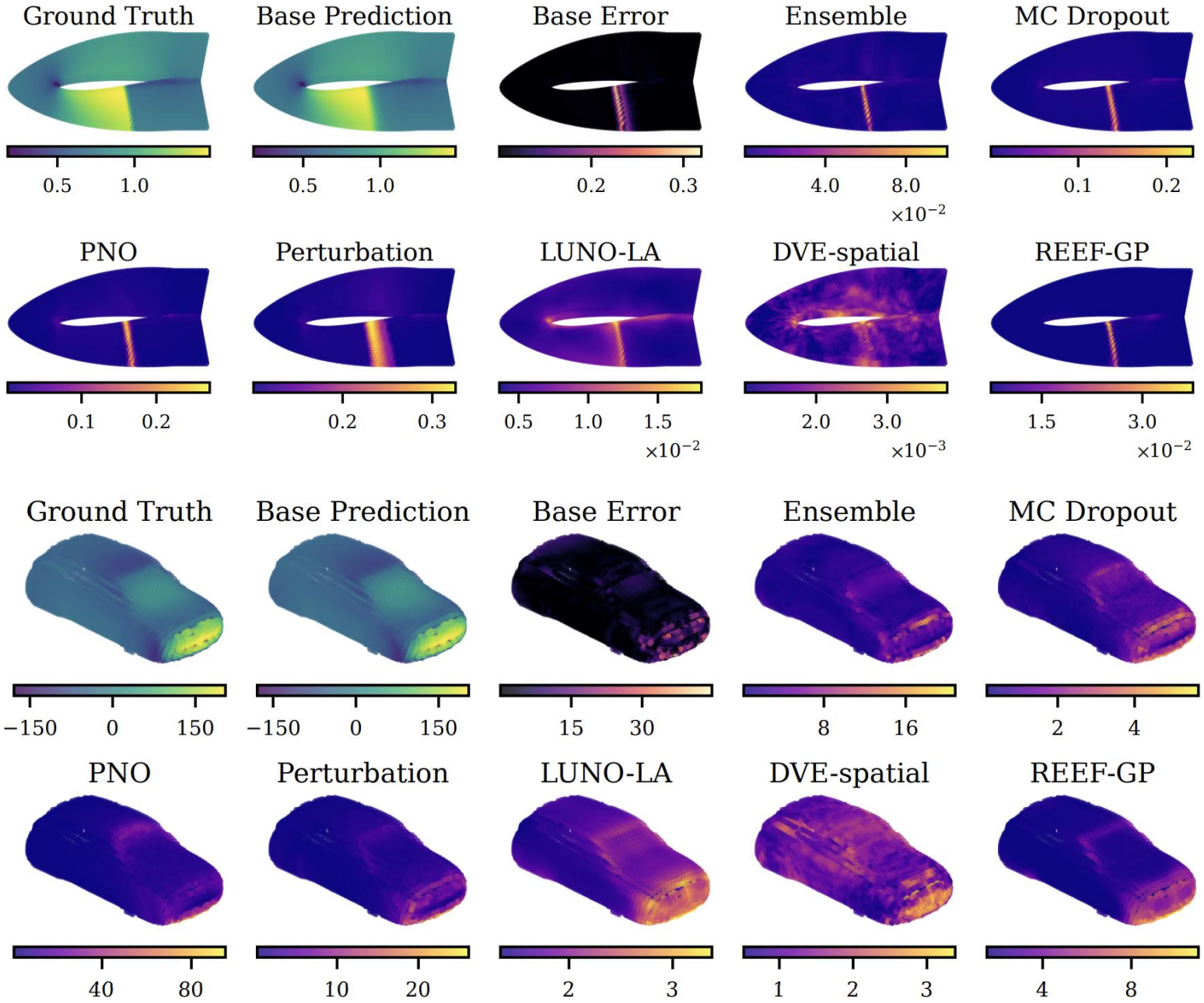
Haixu Wu, Huakun Luo, Haowen Wang, Jianmin Wang, and Mingsheng Long. Transolver: A fast transformer solver for pdes on general geometries, 2024.

RESULTS

Method	rL2 ↓	NLL ↓	CRPS ↓	NIS ↓	ES ↓
Airfoil (2D)	(%)		($\times 10^{-4}$)	($\times 10^{-2}$)	
Base (no UQ)	1.24 ±0.12	–	–	–	–
Ensemble	**1.06**	<u>−3.46</u>	**28.63**	**3.89**	**0.35**
MC Dropout	1.58 ±0.21	−3.38 ±0.15	44.88 ±5.66	<u>5.55</u> ±0.91	0.55 ±0.08
PNO	2.32 ±1.40	−3.28 ±0.65	74.27 ±56.12	8.26 ±6.39	0.79 ±0.47
Perturbation	3.33 ±0.19	−3.32 ±0.05	81.05 ±5.67	9.71 ±0.66	1.08 ±0.06
LUNO-LA	<u>1.24</u> ±0.12	−3.07 ±0.82	<u>38.70</u> ±2.50	6.19 ±0.54	0.47 ±0.04
DVE-spatial	2.09 ±0.25	19.98 ±7.32	88.41 ±12.53	25.89 ±3.09	0.92 ±0.10
REEF-GP	<u>1.24</u> ±0.12	**−3.51** ±0.06	39.80 ±2.35	5.69 ±0.47	<u>0.45</u> ±0.05
ShapeNet Car (3D)	(%)				
Base (no UQ)	8.62 ±0.13	–	–	–	–
Ensemble	**7.73**	<u>4.16</u>	**1.88**	<u>30.28</u>	**194.57**
MC Dropout	9.04 ±0.59	17.21 ±4.21	2.52 ±0.26	67.96 ±9.04	267.68 ±17.66
PNO	13.76 ±1.74	4.98 ±1.52	3.53 ±0.50	51.10 ±12.17	337.16 ±43.04
Perturbation	8.62 ±0.13	13.44 ±1.14	2.28 ±0.03	50.82 ±1.10	224.73 ±3.97
LUNO-LA	8.62 ±0.13	33.55 ±23.43	2.47 ±0.13	73.47 ±14.91	271.32 ±12.10
DVE-spatial	13.08 ±0.26	11.22 ±3.03	3.76 ±0.10	96.17 ±10.43	396.70 ±6.79
REEF-GP	<u>8.52</u> ±0.14	**2.84** ±0.09	<u>2.05</u> ±0.03	**28.81** ±1.65	<u>224.31</u> ±4.18

- No accuracy degradation (rL2).
- Competitive to Deep ensembles on UQ metrics.

UNCERTAINTY IS SPATIALLY MEANINGFUL



- Predicted std tracks where the **base error** is larger.
- Higher uncertainty concentrates at **shock fronts** (Airfoil) and the **front bumper** (Car).
- The distribution is **physically interpretable**.

CONCLUSIONS

Takeaways

- The **coordinate-feature** space of a frozen neural operator is enough for calibrated post-hoc UQ.
- REEF-GP **is competitive with Deep Ensembles** across 5 benchmarks at a fraction of the cost, with **no accuracy loss**.
- Uncertainty is **spatially interpretable** and stable under **geometric shift**.

Paper

Code

ACKNOWLEDGEMENT



NSF



ONR

Thank you!

California-Catalonia Program
BALSELLS FELLOWSHIP

Balsells Fellowship

NEURAL INFORMATION PROCESSING SYSTEMS

NeurIPS