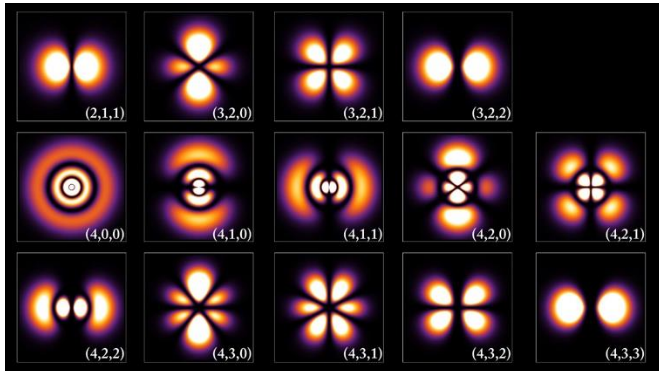
Geometry Aware Operator Transformer as an Efficient and Accurate Neural Surrogate for PDEs on Arbitrary Domains

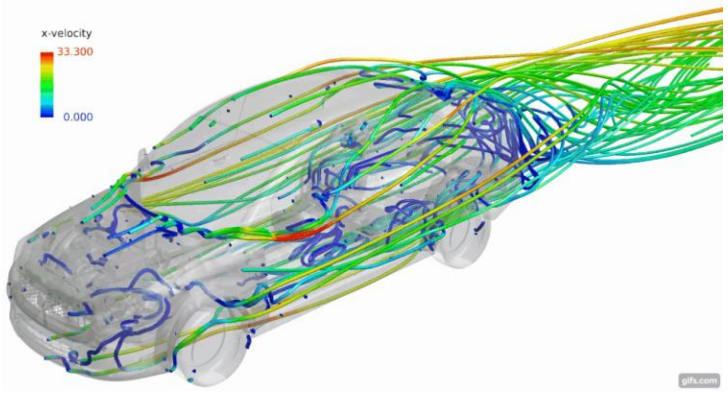
Shizheng Wen[1], Arsh Kumbhat[1], Levi Lingsch[1,2], Sepehr Mousavi[1,3], Yizhou Zhao[4], Praveen Chandrashekar[5], Siddhartha Mishra[1,2]

[1]Seminar for Applied Mathematics, ETH Zurich, Switzerland
[2]ETH AI Center, Zurich, Switzerland
[3]Department of Mechanical and Process Engineering, ETH Zurich, Switzerland
[4]School of Computer Science, CMU, USA
[5]Centre for Applicable Mathematics, TIFR, India

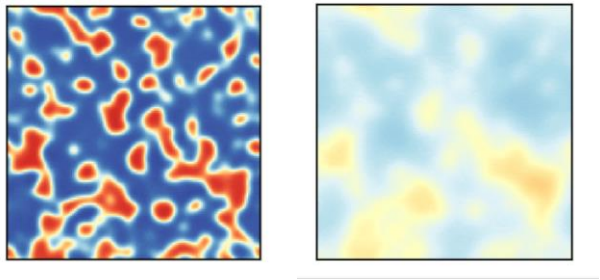
PDE are building blocks of science



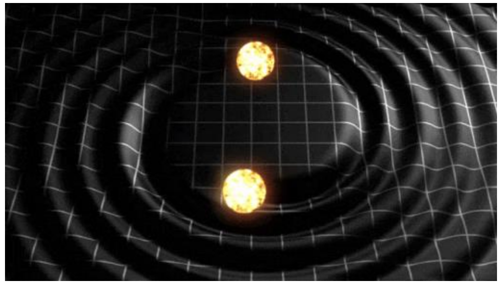
Schrödinger equation



Navier-Stokes equations

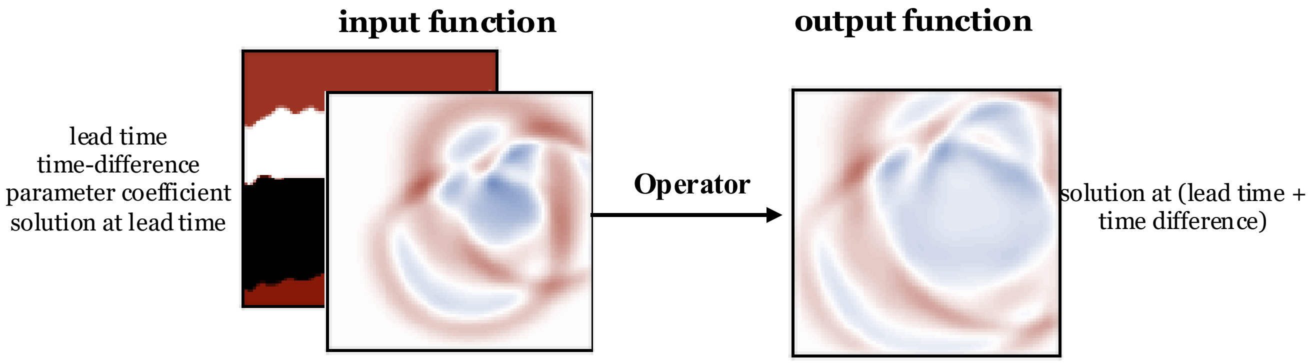


Reactions-Diffusion equations



Einstein field equations

Background – Learn Solution Operator from Data



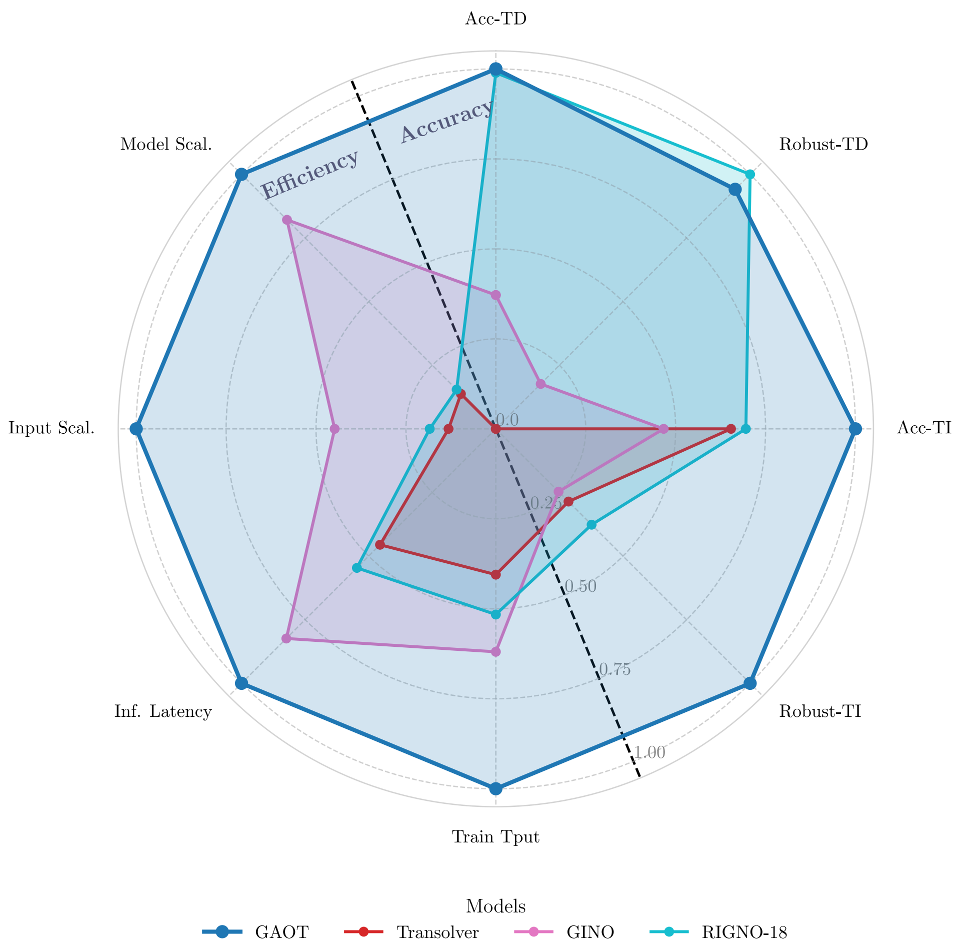
Time-Shifted Operator:

$$\mathcal{S}^{\dagger} : \mathcal{X} \times \mathcal{Q} \times (0,T) \times \mathbb{R}^{+} \longrightarrow \mathcal{X}$$

$$\mathcal{S}^{\dagger}\big(u^t,\, c^t,\, t,\, \tau\big) \;=\; \mathcal{S}^{t}\big(u^t,\, c^t,\, \tau\big) \;=\; u^{t+\tau}$$

- Use machine learning to directly learn the solution operator from data.

Motivation – Efficient and Accurate Neural Solver on Arbitrary domains



Neural Solver on arbitrary domains

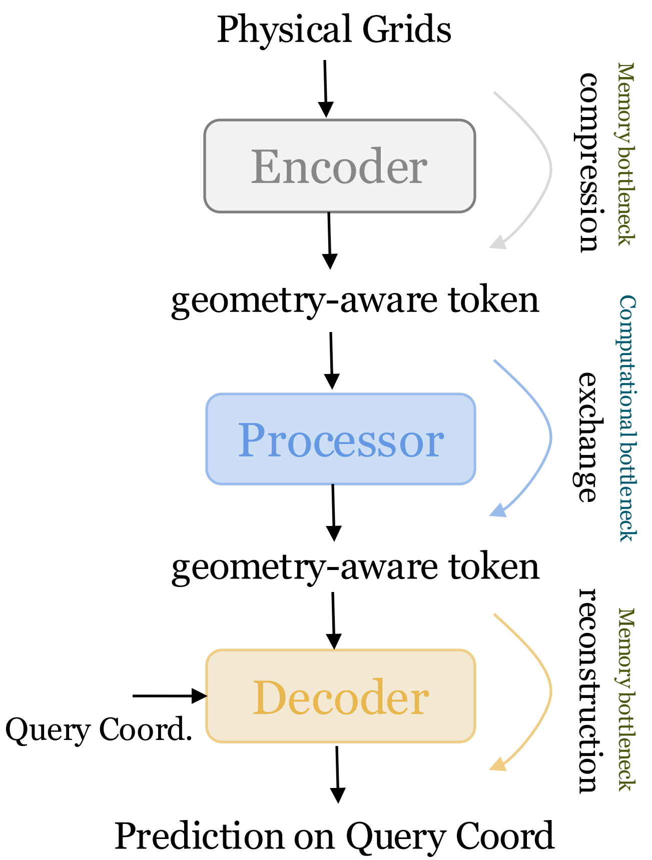
Accuracy-Efficiency Tradeoff

- Accuracy and robust models, such as graph-based ones (RIGNO), are not necessarily computational efficient nor scalable.
- Efficient models, such as Fourier-based ones (GINO), are not accurate enough.

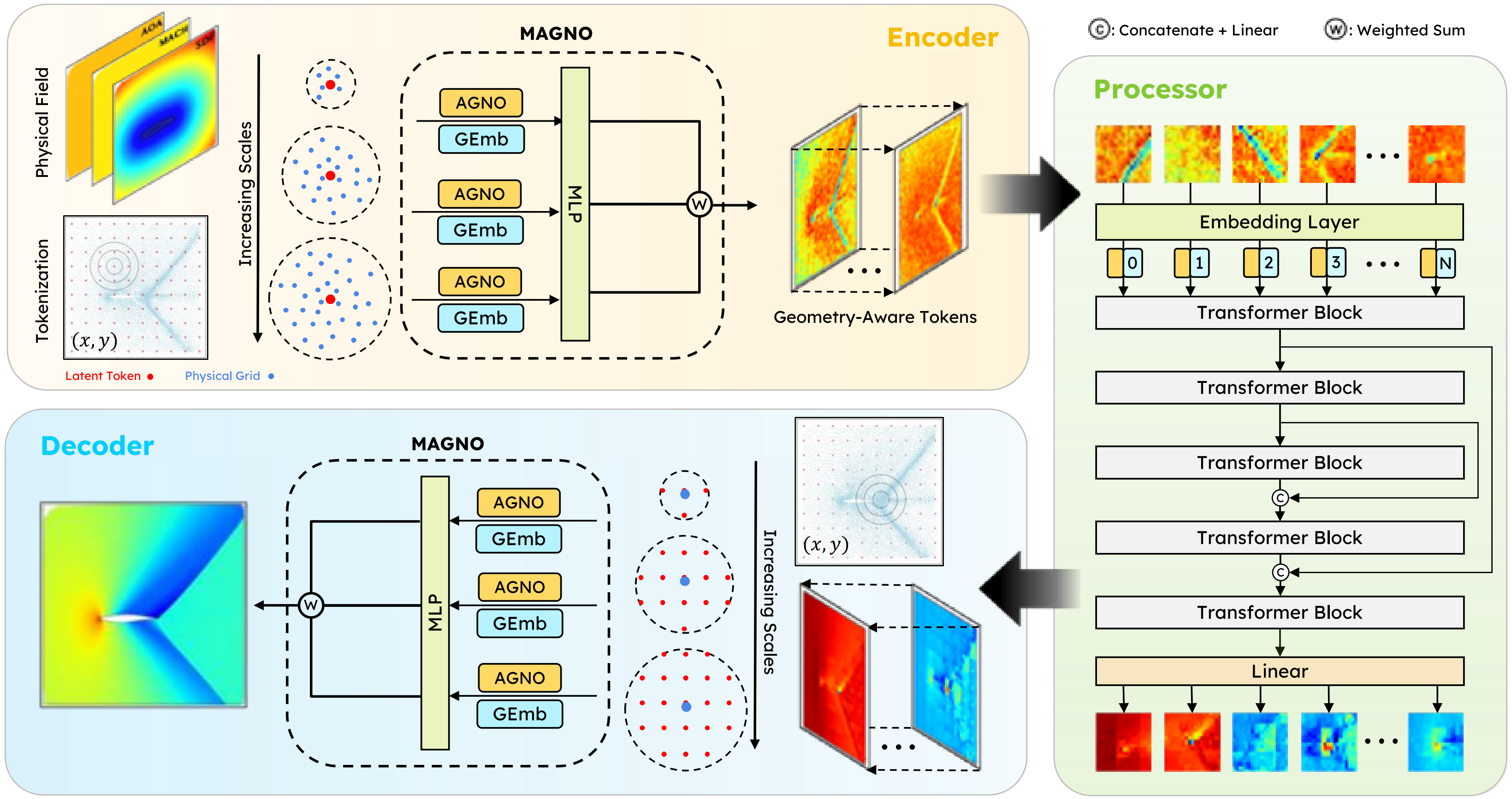
Geometry Aware Operator Transformer (GAOT)

- Scalable in model size and input size
- High train throughput and inference latency
- Accurate and robust predictions in TD and TI.

Overview of GAOT



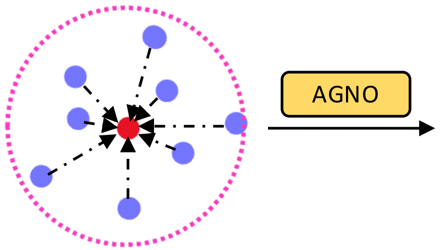
Geometry Aware Operator Transformer



Encoder/Decoder: Local + Multiscale information

Processor: Fusion of global information/different modal tokens

Accuracy Design Highlight - AGNO

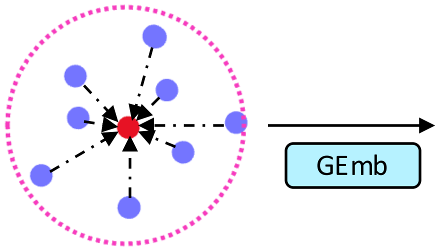


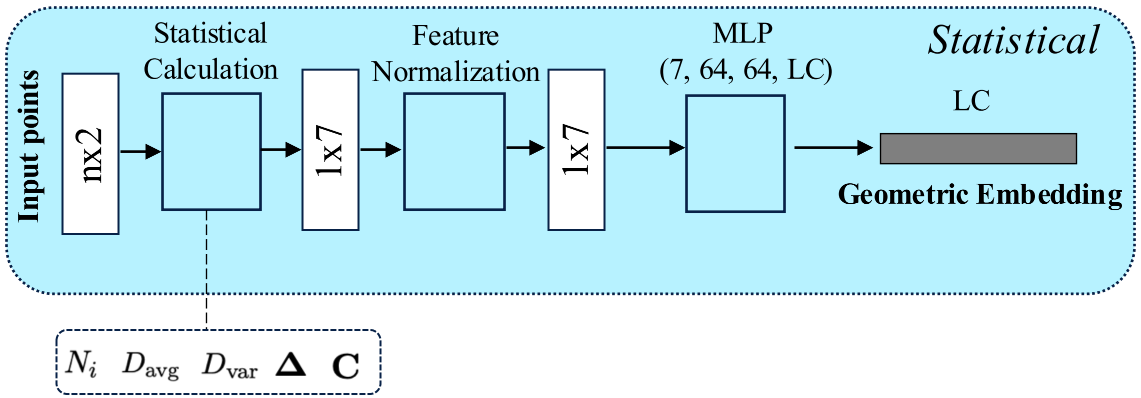
Trick 1: Attentional Weighting in Local Integration

$$(\mathcal{L}_r f)(x) = \int_{A_r(x)} K_\ell\big(x, y, f(y)\big)\, \varphi\big(f(y)\big)\, \mathrm{d}y \approx \sum_{i=1}^{n_y} \boxed{\alpha_i}\, K_\ell\big(x, y_i, f(y_i)\big)\, \varphi\big(f(y_i)\big)$$

$$\alpha_i = \frac{\exp(e_i)}{\sum_{j=1}^{n_y} \exp(e_j)} \qquad e_i = \frac{\langle W_q\, \mathbf{x},\ W_k\, \mathbf{y}_i \rangle}{\sqrt{d}}$$

Accuracy Design Highlight - GEmb



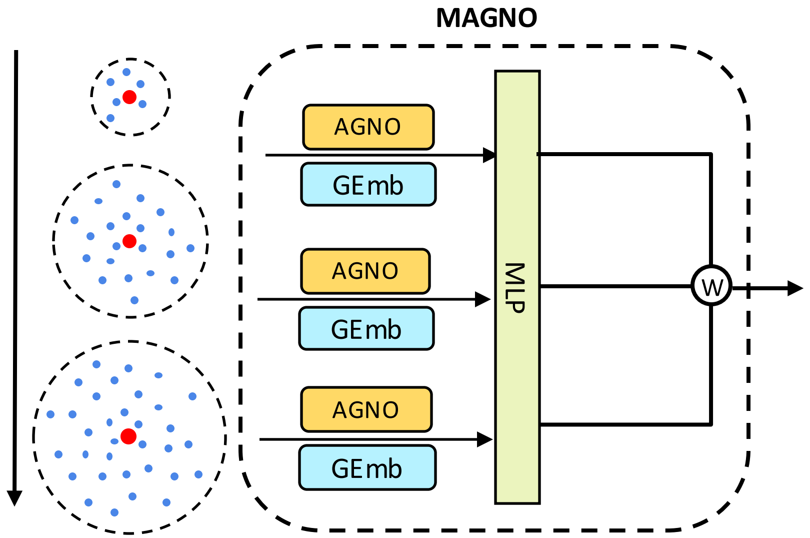


Trick 2: Lightweight Geometric Embedding

Statistical descriptors

1. Number of neighbors
2. Average Distance
3. Distance Variance
4. Centroid Offset Vector
5. PCA Features

Accuracy Design Highlight - MAGNO



Trick 3: Multiscale Neighborhood Construction + self-adaptive weighting sum

$$A_{r_1}(x), \quad A_{r_2}(x), \quad \ldots, \quad A_{r_M}(x).$$

$$\big(\mathcal{L}_{\text{multi}} f\big)(x) \;=\; \sum_{m=1}^{M} \boxed{\beta_m(x)}\, \big(\mathcal{L}_{r_m} f\big)(x).$$

$$\boldsymbol{g}(x) \;=\; \mathrm{MLP}_{\alpha}(\mathbf{x}) \;\in\; \mathbb{R}^{M}$$

$$\beta_m(x) \;=\; \frac{\exp\!\big(g_m(x)\big)}{\sum_{m'=1}^{M} \exp\!\big(g_{m'}(x)\big)}, \qquad m = 1, \ldots, M.$$

- Introduce learning components in determining the importance of different scales features.
- Analogy of the different size of kernels in convolutional neural network.

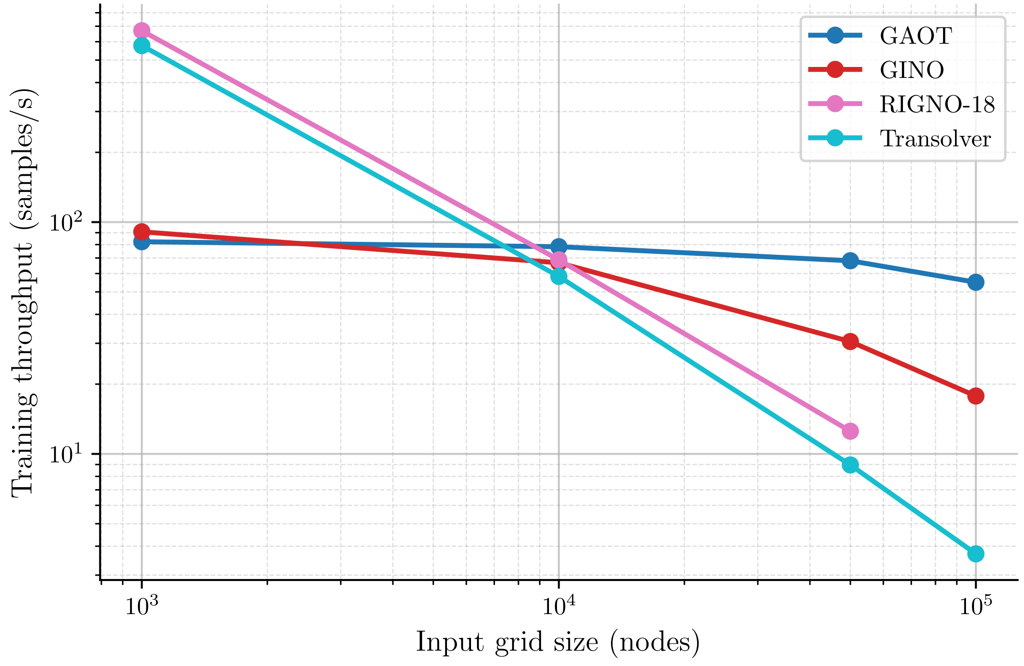
Efficiency Design Highlights

- **Trick 1: Computation and Memory Allocation:** Lightweight graph-based encoder/decoder (sparse operations) shifts most computation to the structured-grid transformer processor (computational-intensive), improving throughput. Sequential process every data in each batch for Encoder and Decoder to avoid OOM, and batch processing in Processor for improving Tput.

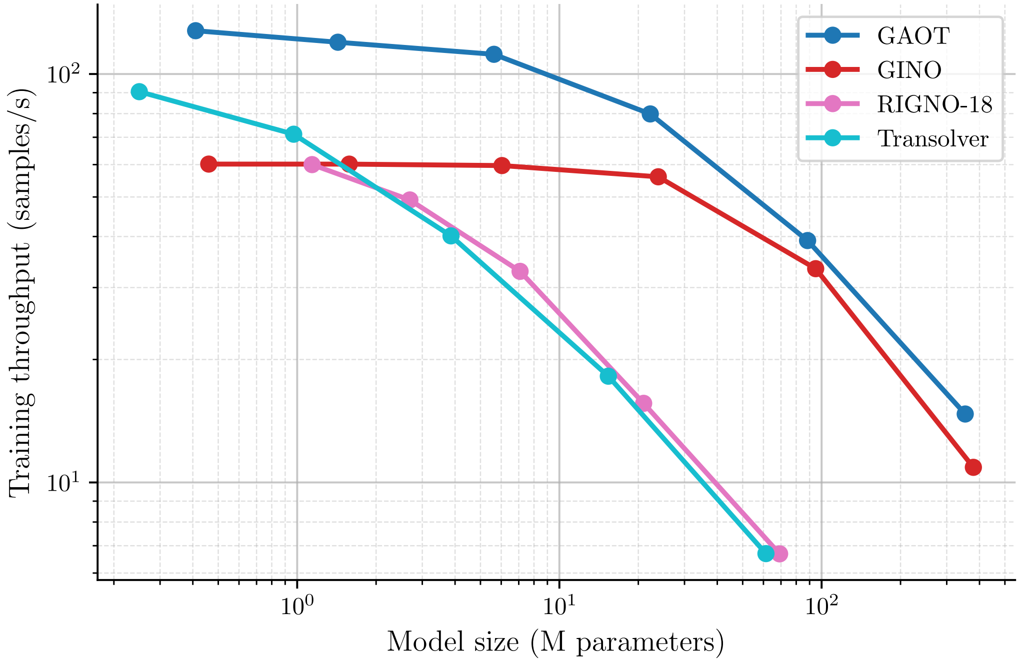
- **Trick 2: Memory-Efficient Graph Construction:** Edge Masking reduces graph density, alleviating memory bottlenecks and enhancing generalization.

- **Trick 3: Graph Cache or Asynchronous Pipeline:** Graph caching or asynchronous pipeline in building the graph in the data-loader to avoid redundant graph construction and improve training efficiency.

Scalable Experiments



Grid vs. Throughput



MODEL vs. Throughput

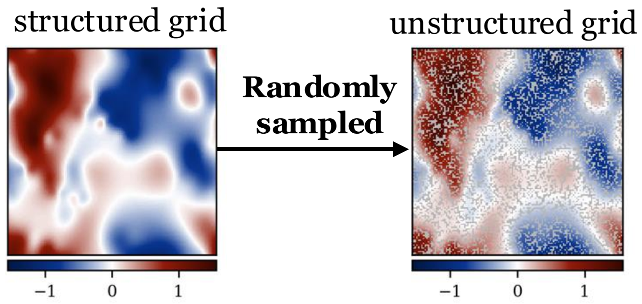
- GAOT scales much more favorably than the baselines with respect to both input and model size.

Datasets

Abbreviation	GeoVar	Characteristic	PDE Type
Poisson-C-Sines*	F	Circular domain with sines f	PE
Poisson-Gauss	F	Gaussian source	PE
Elasticity	T	Hole boundary distance	HEE
NACA0012*	T	Flow past NACA0012 airfoil	CE
NACA2412*	T	Flow past NACA2412 airfoil	CE
RAE2822*	T	Flow past RAE2822 airfoil	CE
Bluff-Body*	T	Flow past bluff-bodies	CE
DrivAerNet++(p)	T	Surface pressure	INS
DrivAerNet++(wss)	T	Surface wall shear stress	INS
DrivAerML(p)	T	Surface pressure coefficient	INS
DrivAerML(wss)	T	Surface wall shear stress	INS
NASA-CRM (p)	T	Surface pressure	INS
NASA-CRM (sfc)	T	Surface friction coefficient	INS
NS-Gauss	F	Gaussian vorticity IC	INS
NS-PwC	F	Piecewise const. IC	INS
NS-SL	F	Shear layer IC	INS
NS-SVS	F	Sinusoidal vortex sheet IC	INS
CE-Gauss	F	Gaussian vorticity IC	CE
CE-RP	F	4-quadrant RP	CE
Wave-Layer	F	Layered wave medium	WE
Wave-C-Sines	F	Circular domain with sines IC	WE

Data Source:

- Public structured Dataset (PDEgym) and randomly sampled ones.
- Public unstructured dataset (Elasticity, Wave-C-Sines, DrivAerML, NASA CRM, DrivAerNet++)
- Self-generated unstructured dataset (Poisson-C-Sines, NACA0012, NACA2412, RAE2822, Bluff-Body)



Data Type:

- Compressible Euler, Incompressible NS
- Poisson Equation
- Wave Equation
- Hyper-elastic equation

- Extensively testing GAOT on 28 challenging benchmarks for both time-independent and time-dependent PDEs of various types, ranging from regular grids to random point clouds to highly unstructured adapted grids.

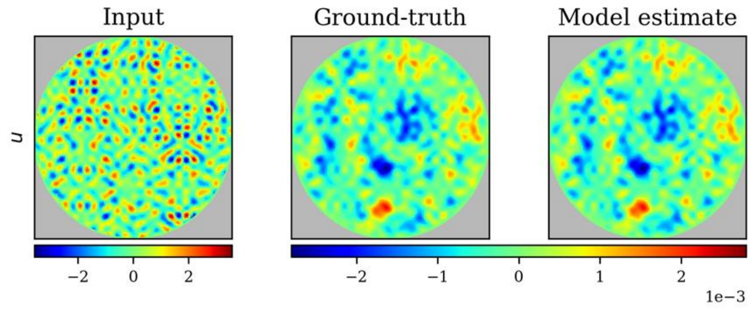
Overall Result – unstructured grid dataset

Table 1: Benchmark results on time-dependent and time-independent datasets. Best and 2nd best models are shown in blue and orange fonts for each dataset.

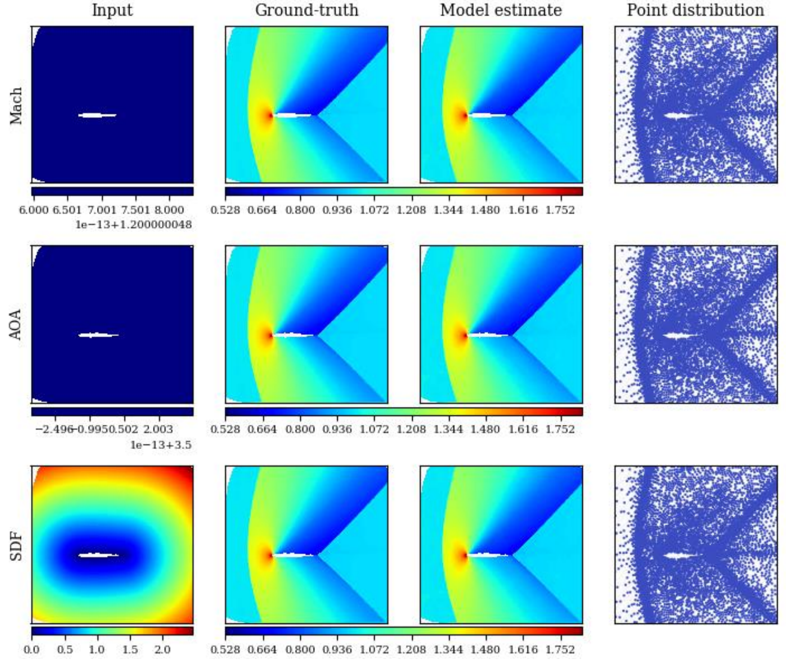
Dataset	Median relative L^1 error [%]					
Time-Independent	**GAOT**	**RIGNO-18**	**Transolver**	**GNOT**	**UPT**	**GINO**
Poisson-C-Sines	3.10	6.83	77.3	100	100	20.0
Poisson-Gauss	0.83	2.26	2.02	88.9	48.4	7.57
Elasticity	1.34	4.31	4.92	10.4	12.6	4.38
NACA0012	6.81	5.30	8.69	6.89	16.1	9.01
NACA2412	6.66	6.72	8.51	8.82	17.9	9.39
RAE2822	6.61	5.06	4.82	7.15	16.1	8.61
Bluff-Body	2.25	5.76	1.78	44.2	5.81	3.49
Time-Dependent	**GAOT**	**RIGNO-18**	**GeoFNO**	**FNO DSE**	**UPT**	**GINO**
NS-Gauss	2.91	2.29	41.1	38.4	92.5	13.1
NS-PwC	1.50	1.58	26.0	56.7	100	5.85
NS-SL	1.21	1.28	13.7	22.6	51.5	4.48
NS-SVS	0.46	0.56	9.75	26.0	4.2	1.19
CE-Gauss	6.40	6.90	42.1	30.8	64.2	25.1
CE-RP	5.97	3.98	18.4	27.7	26.8	12.3
Wave-Layer	5.78	6.77	11.1	28.3	19.6	19.2
Wave-C-Sines	4.65	5.35	13.1	5.52	12.7	5.82

- GAOT is very accurate on all of them, being either the best (10) or second-best (4) model on 14 of them.

Visualization Examples



Poisson-C-Sines



NACA0012

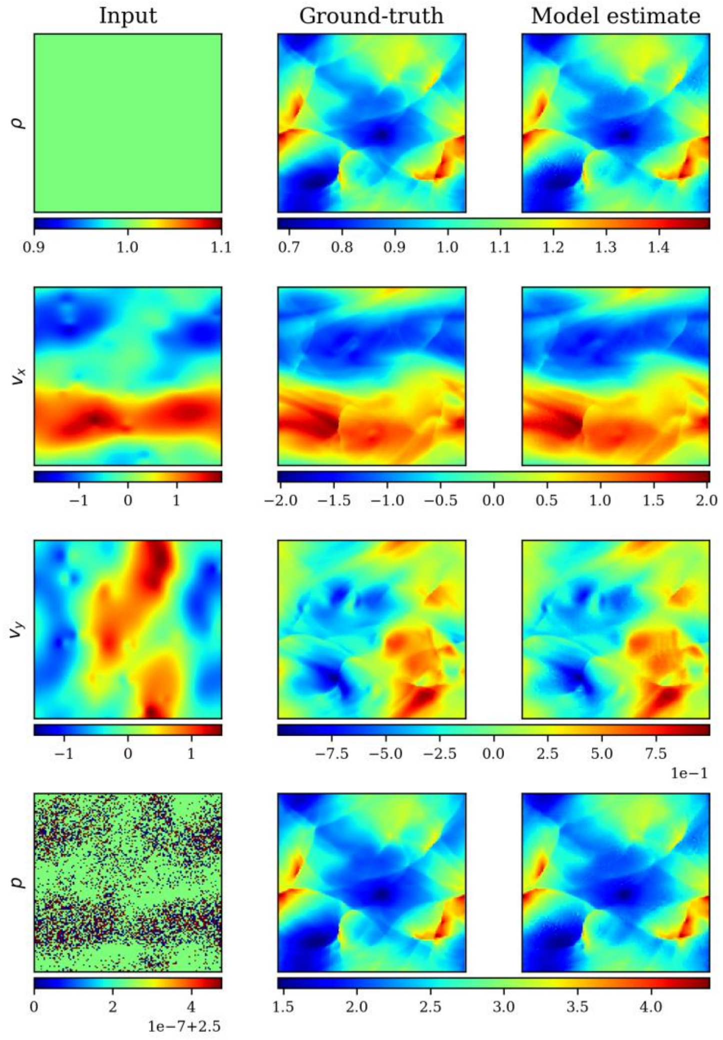
Overall Result – structured grid dataset

Table E.5: Benchmarks with time-dependent datasets with regular grid inputs. Best and 2nd best models are shown in blue and orange fonts for each dataset.

Dataset	Median relative L^1 error [%]						
Structured	**GAOT**	**RIGNO-18**	**RIGNO-12**	**CNO**	**ViT**	**scOT**	**FNO**
NS-Gauss	2.29	2.74	3.78	10.9	3.16	2.92	14.41
NS-PwC	1.23	1.12	1.82	5.03	3.89	7.12	12.55
NS-SL	0.98	1.13	1.82	2.12	0.73	2.49	2.08
NS-SVS	0.46	0.56	0.75	0.70	0.39	1.01	7.52
CE-Gauss	5.28	5.47	7.56	22.0	6.81	9.44	28.69
CE-RP	4.98	3.49	4.43	18.4	4.30	9.74	38.48
Wave-Layer	5.40	6.75	8.97	8.28	5.48	13.44	28.13

- GAOT consistently ranks within the top two across six of the seven benchmark datasets

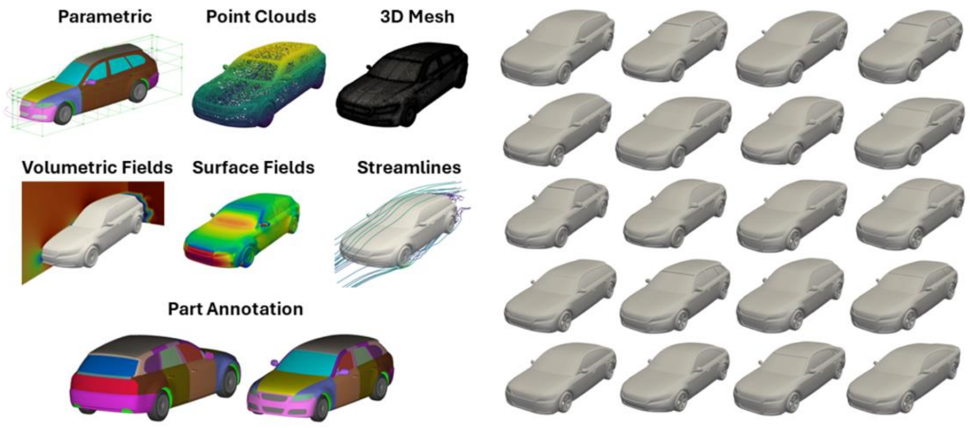
Visualization Examples

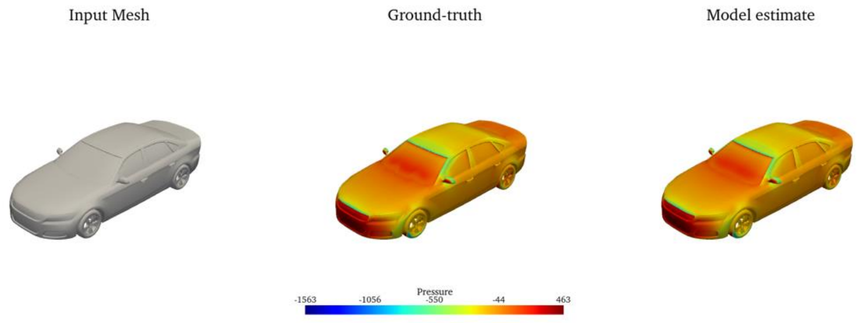


CE-Gauss

Overall Result – DrivAerNet++

- Large Scale Flow past Cars Dataset (8K Car Shapes)
- 500K Surface nodes

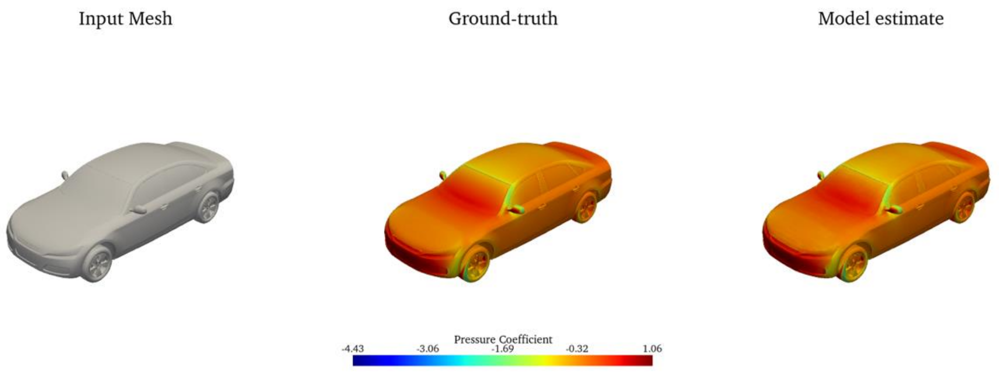


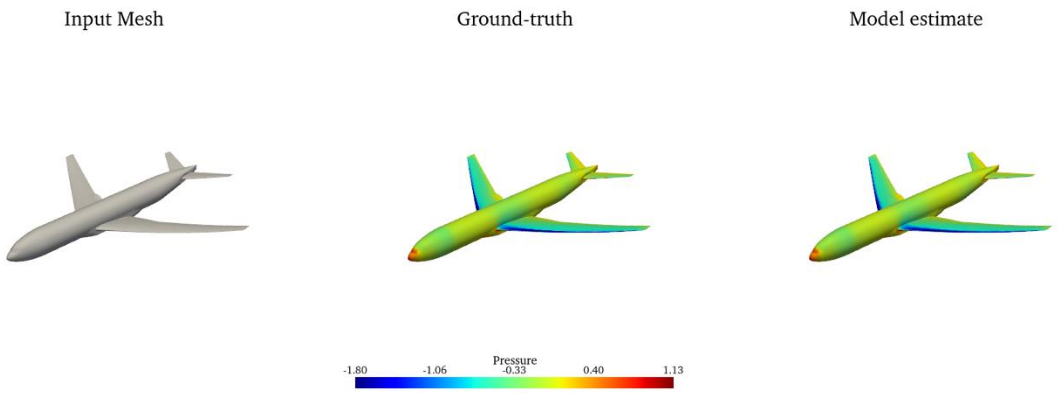


Model	Pressure		Wall Shear Stress	
	MSE	Mean AE	MSE	Mean AE
GAOT	4.2694	1.0699	8.6878	1.5429
FIGConvNet	4.9900	1.2200	9.8600	2.2200
TripNet	5.1400	1.2500	9.5200	2.1500
RegDGCNN	8.2900	1.6100	13.8200	3.6400
GAOT (NeurField)	12.0786	1.7826	22.9160	2.5099

- GAOT: SOTA for Surface Pressure, Shear Stress

Overall Result – DrivAerML and NASA CRM





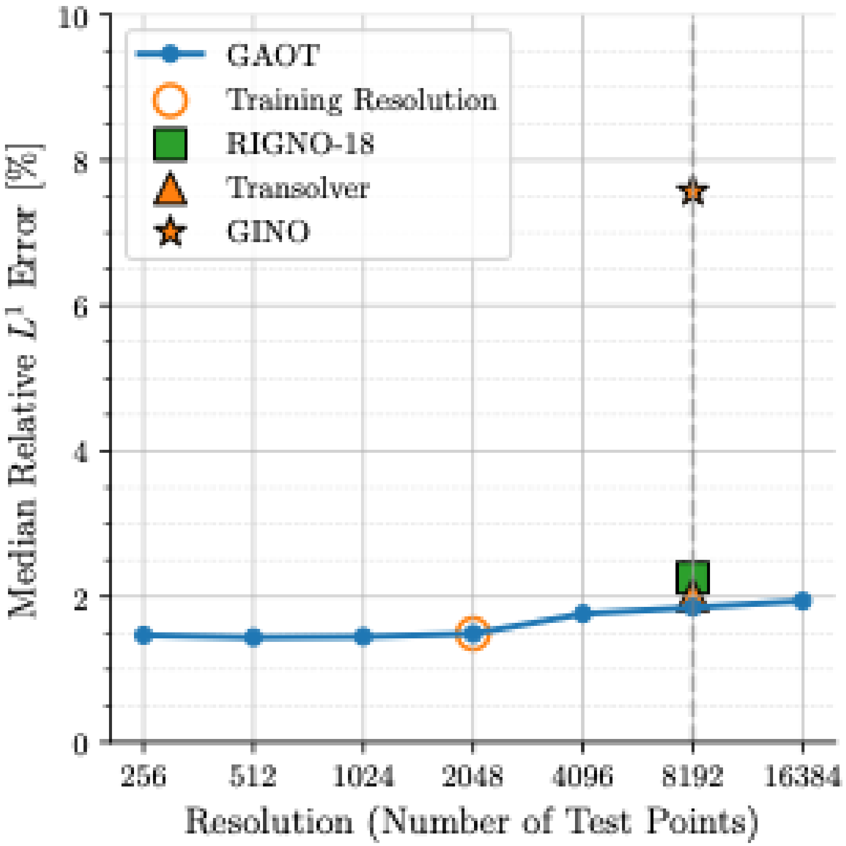
- HR-LES simulations of flow past 500 cars.
- More accurate than RANS for DrivAerNet++
- Up to 10M surface nodes!

- AIAA's NASA CRM Benchmark
- 500K surface nodes

Model	Cp (DML)		WSS (DML)		P (CRM)		Cf (CRM)	
	MSE	Mean AE	MSE	Mean AE	MSE	Mean AE	MSE	Mean AE
GAOT	5.1729	1.2352	16.9818	2.1640	7.7170	1.6014	16.1091	2.2305
GINO	8.8124	1.5238	28.4832	2.7330	10.5688	1.7450	21.1789	2.4240

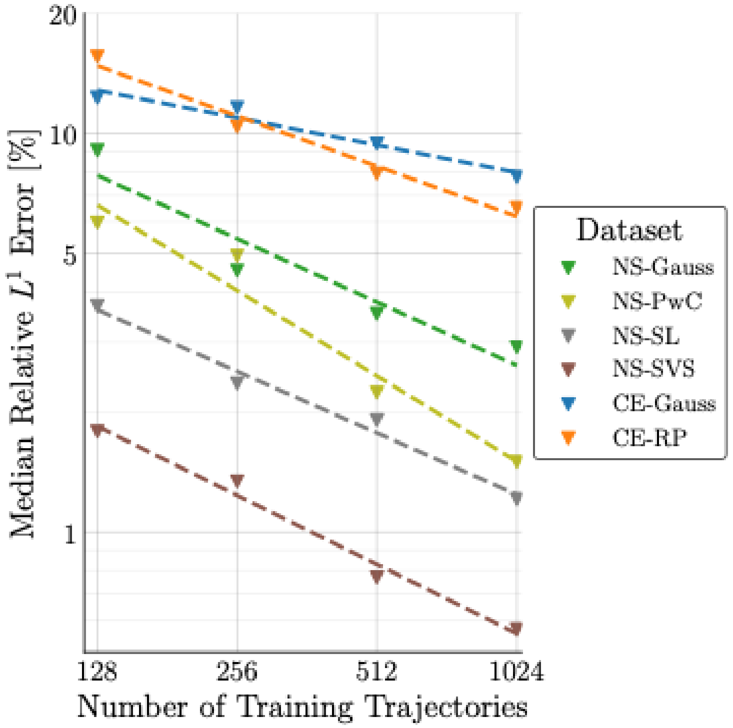
- GATO accurately predicts surface pressure coefficient + wall shear stress for DrivAerML and surface pressure + skin friction for NASA CRM.

Overall Result – Resolution invariance

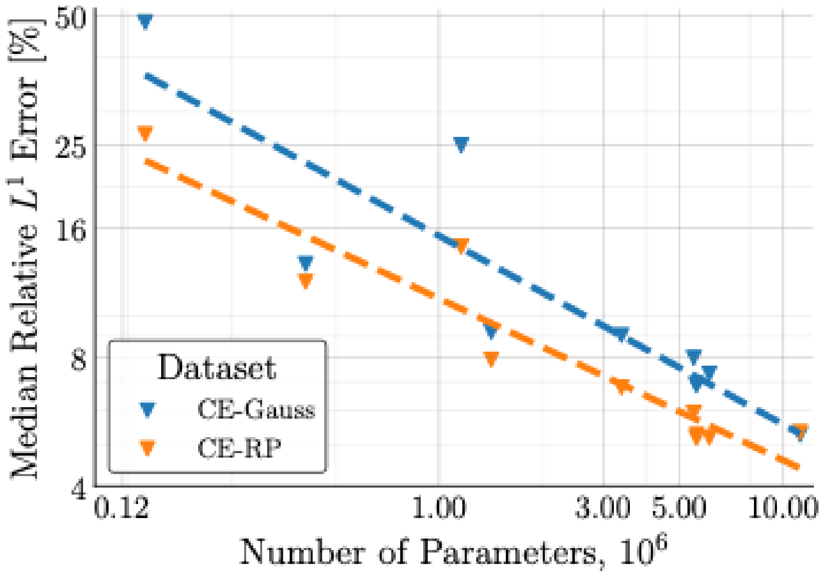


- GAOT model is trained at a resolution of 2048 and evaluate at various test resolutions. The results for RIGNO-18, Transolver and GINO correspond to models trained and tested at a resolution of 8192.
- GAOT possesses excellent resolution invariance. It also achieves the best performance when tested at 8192 points.

Overall Results – Scaling with model size and dataset size



Scaling Data Size

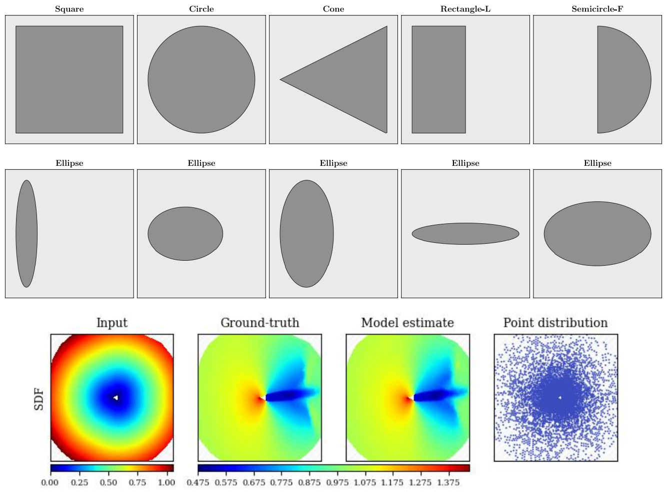


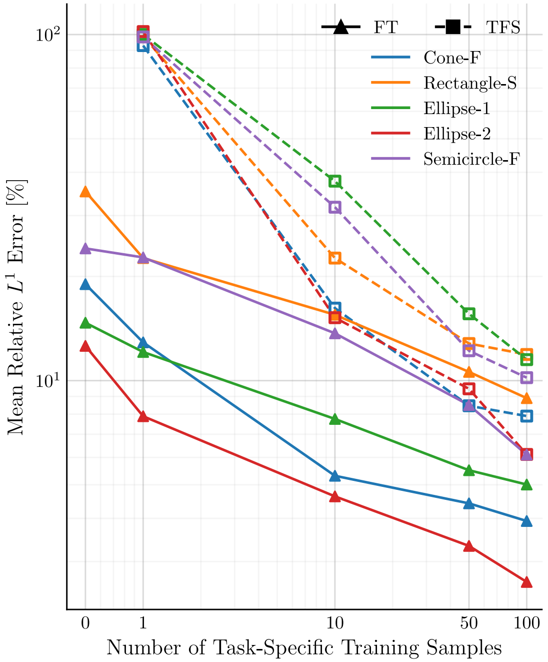
Scaling Model Size

- GAOT shows excellent scaling law with the model size and dataset size.

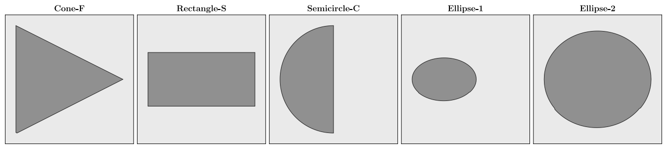
Overall Result – Transfer Learning

Pretraining Shapes

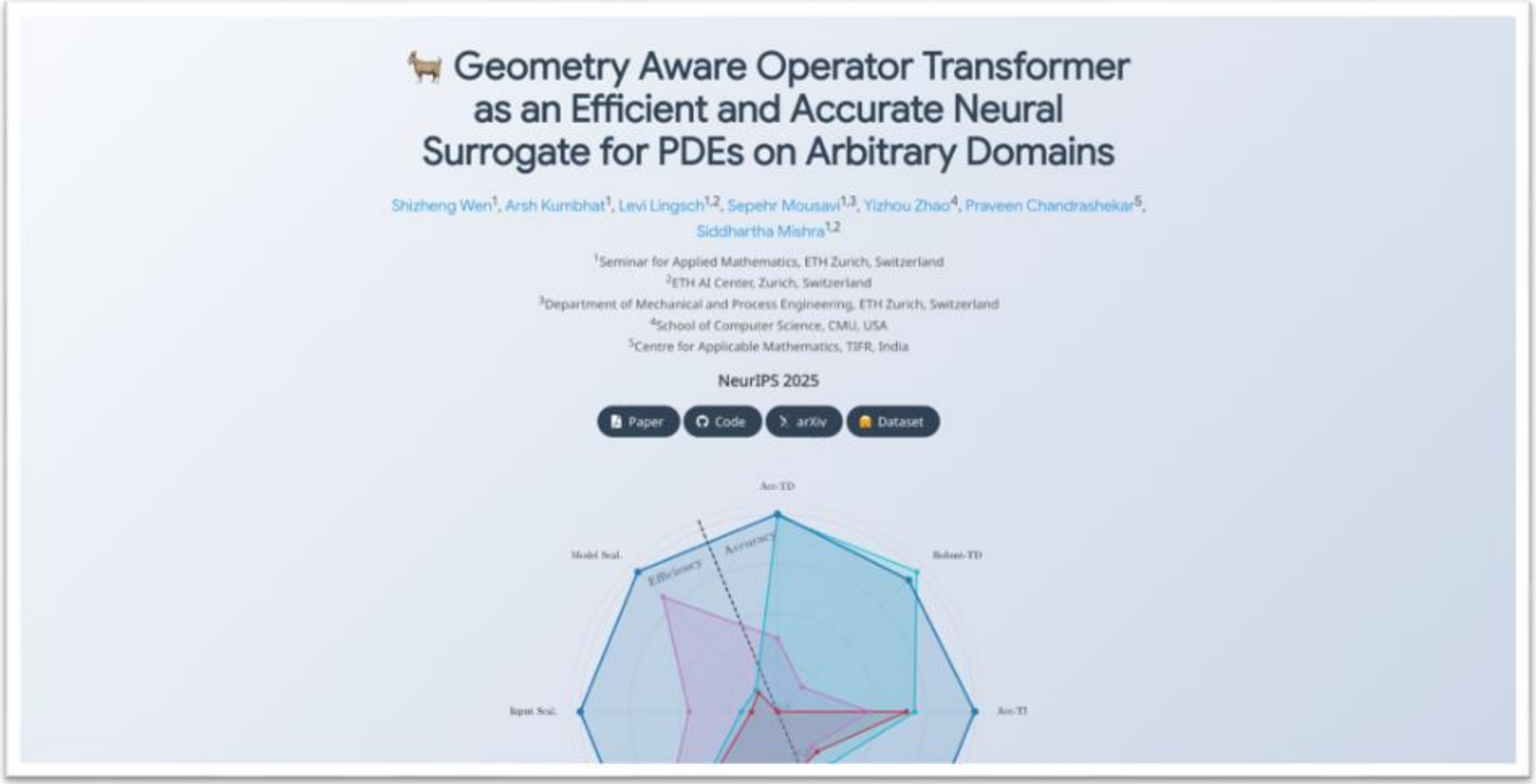




Finetuning Shapes



- GAOT performs very well in few-shot transfer learning scenario.
- Fine-tuned model (FT) providing an almost order of magnitude gain in accuracy over the model, trained from scratch (TFS)



Project Page: https://camlab-ethz.github.io/GAOT/