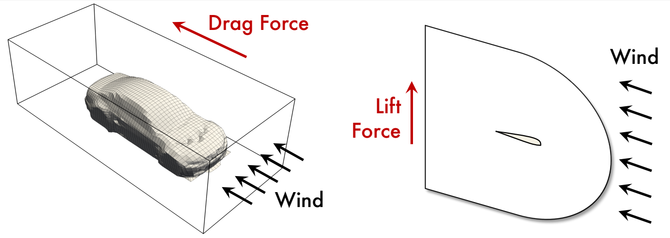
Transolver: A Fast Transformer Solver for PDEs on General Geometries

Tsinghua University

Haixu Wu, Huakun Luo, Haowen Wang, Jianmin Wang, Mingsheng Long#

ICML | 2024
Spotlight

A Practical PDE Solver Should Handle Large-Scale Unstructured Meshes



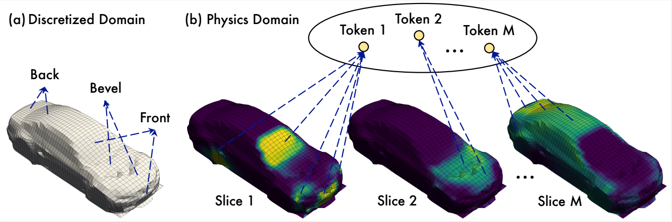
(a) Shape-Net Car (b) AirfRANS

Previous work: Geometric deep learning, Linear Transformer (directly applying attention to mesh points, **over 10k tokens**)

RULER: The effective length of GPT-4 is 64k

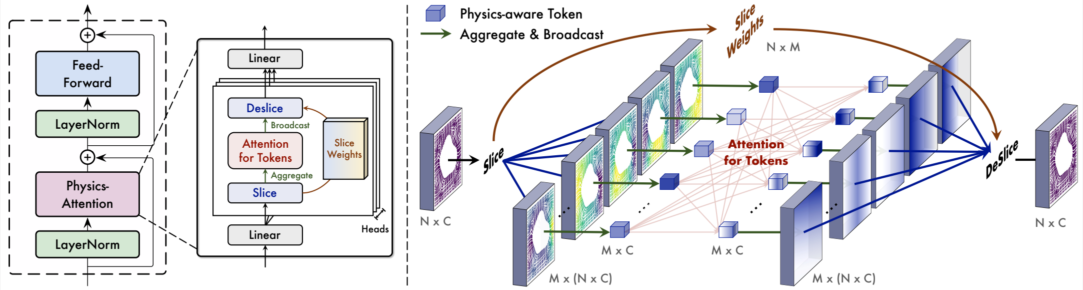
Challenges: Geometry, Physics, Efficiency

Key Insight: Getting rid of superficial and unwieldy meshes, learning **intrinsic physical states** underlying geometries



All the code, data and scripts are in
https://github.com/thuml/Transolver

Transolver: Making Transformers good at solving PDEs



➢ **Tokenize:** Learning physics-aware tokens

$$\{\mathbf{w}_i\}_{i=1}^N = \{\mathrm{Softmax}\left(\mathrm{Project}\left(\mathbf{x}_i\right)\right)\}_{i=1}^N$$

$$\mathbf{s}_j = \{\mathbf{w}_{i,j}\mathbf{x}_i\}_{i=1}^N,$$

$$\mathbf{z}_j = \frac{\sum_{i=1}^N \mathbf{s}_{j,i}}{\sum_{i=1}^N \mathbf{w}_{i,j}} = \frac{\sum_{i=1}^N \mathbf{w}_{i,j}\mathbf{x}_i}{\sum_{i=1}^N \mathbf{w}_{i,j}}$$

① Assign each point to M "classes"

② Separate feature to M "slices"

③ Globally weighted sum features on each slice for M physics-aware tokens

➢ **Physics-Attention:** Applying attention to learned physics tokens

Physics-Attention is equivalent to learnable integral on input domain

22% Error Reduction on Six Standard Benchmarks
Excels in Industrial-level simulations

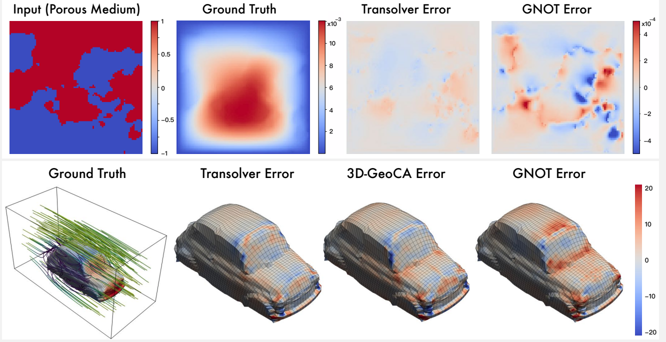
➢ **More than 20 baselines:** FNO, LSM, GNOT, ONO, GUNet, etc

➢ **Number of Mesh points:** ranging from 972 to 32,186

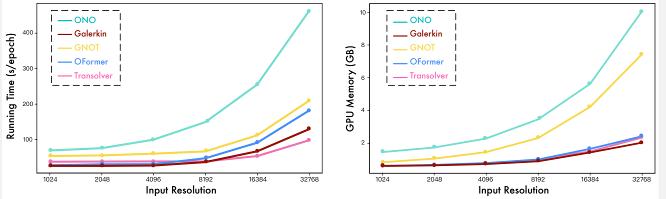
Model ($\times 10^{-2}$)	Point Cloud	Structured Mesh			Regular Grid		Unstructured Mesh	
	Elasticity	Plasticity	Airfoil	Pipe	Navier–Stokes	Darcy	Shape-Net Car	AirfRANS
Previous SOTA	0.86 (GNOT)	0.17 (OFormer)	0.59 (LSM)	0.47 (GNOT)	11.95 (ONO)	0.65 (LSM)	98.42 (3D-GeoCA)	99.64 (GraphSAGE)
Transolver Promotion	**0.64** 25.6%	**0.12** 29.4%	**0.53** 10.2%	**0.33** 29.7%	**9.00** 24.7%	**0.57** 12.3%	**99.35** -	**99.78** -

We list Spearman's correlation of drag/lift coefficient for Car and AirfRANS, rL2 for others

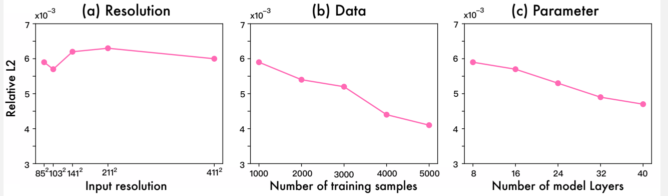
Showcases: Complex geometries, hybrid physics



Efficiency: Linear complexity w.r.t mesh points



Scalability: Benefit from large data and model



OOD generalizability: Test on unseen PDEs

Models	OOD Reynolds		OOD Angles	
	$C_L \downarrow$	$\rho_L \uparrow$	$C_L \downarrow$	$\rho_L \uparrow$
Simple MLP	0.6205	0.9578	0.4128	0.9572
GraphSAGE (2017)	0.4333	0.9707	<u>0.2538</u>	0.9894
PointNet (2017)	0.3836	0.9806	0.4425	0.9784
Graph U-Net (2019)	0.4664	0.9645	0.3756	0.9816
MeshGraphNet (2021)	1.7718	0.7631	0.6525	0.8927
GNO (2020a)	0.4408	<u>0.9878</u>	0.3038	0.9884
Galerkin (2021)	0.4615	0.9826	0.3814	0.9821
GNOT (2023)	<u>0.3268</u>	0.9865	0.3497	0.9868
GINO (2023a)	0.4180	0.9645	0.2583	<u>0.9923</u>
Transolver (Ours)	**0.2996**	**0.9896**	**0.1500**	**0.9950**