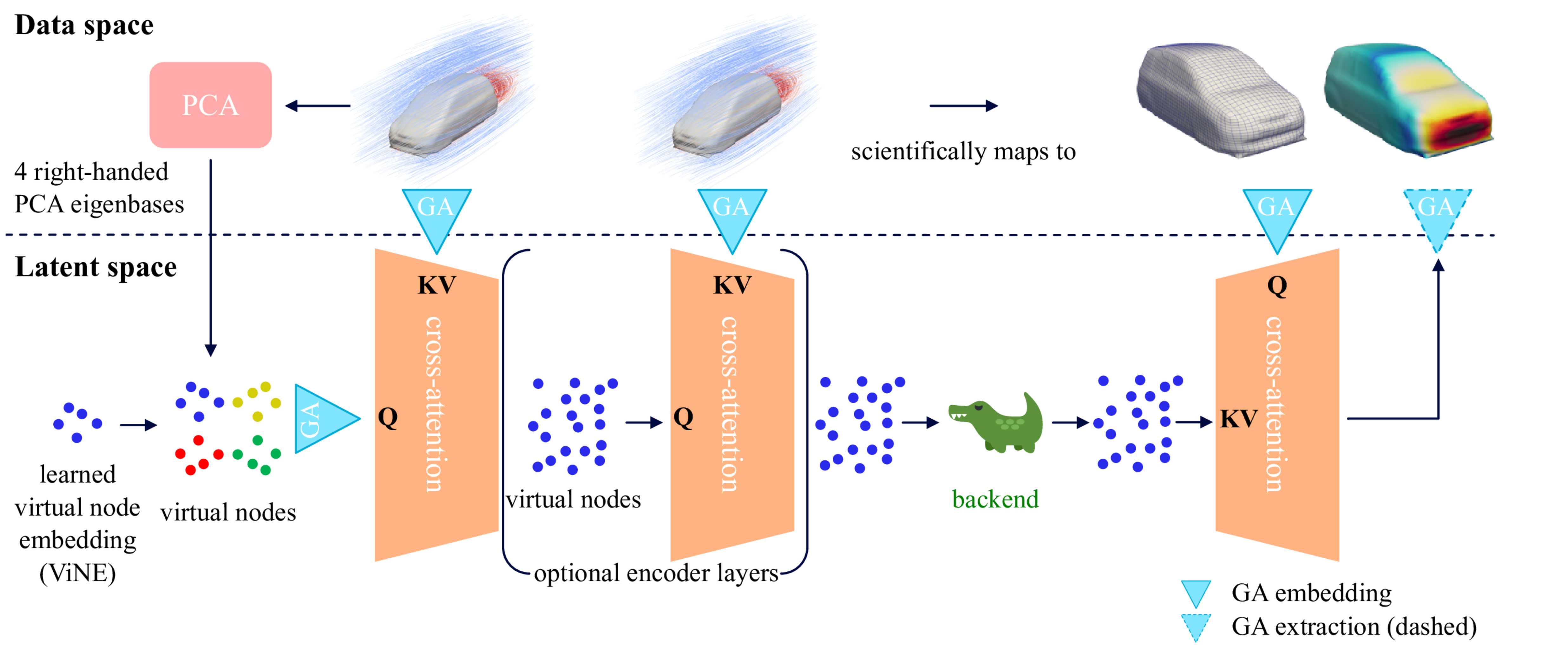
ViNE-GATr: Scaling Geometric Algebra Transformers with Virtual Nodes Embeddings

Julian Suk*, Thomas Hehn, Arash Behboodi, Gabriele Cesa

*Work done during the internship at Qualcomm AI Research, Amsterdam.

Qualcomm AI research

ICLR



Geometric Algebra Transformer

+

PerceiverIO

=

Scalable equivariant transformer efficiently processing geometric data in coarse-grained latent

Problem

Large-scale geometric data typical in many scientific disciplines but requires efficient solutions

Typically, coarsening via *downsampling disrupts the symmetries*

Projective Geometric Algebra and GATr

Geometric Algebra Transformers (GATr) are $E(3)$ equivariant architectures

Object / operator	Scalar 1	Vector e_0	Vector e_i	Bivector e_{0i}	Bivector e_{ij}	Trivector e_{0ij}	Trivector e_{123}	PS e_{0123}
Scalar $\lambda \in \mathbb{R}$	λ	0	0	0	0	0	0	0
Plane w/ normal $n \in \mathbb{R}^3$, origin shift $d \in \mathbb{R}$	0	d	n	0	0	0	0	0
Line w/ direction $n \in \mathbb{R}^3$, orthogonal shift $s \in \mathbb{R}^3$	0	0	0	s	n	0	0	0
Point $p \in \mathbb{R}^3$	0	0	0	0	0	p	1	0
Pseudoscalar $\mu \in \mathbb{R}$	0	0	0	0	0	0	0	μ
Reflection through plane w/ normal $n \in \mathbb{R}^3$, origin shift $d \in \mathbb{R}$	0	d	n	0	0	0	0	0
Translation $t \in \mathbb{R}^3$	1	0	0	$\frac{1}{2}t$	0	0	0	0
Rotation expressed as quaternion $q \in \mathbb{R}^4$	q_0	0	0	0	q_i	0	0	0
Point reflection through $p \in \mathbb{R}^3$	0	0	0	0	0	p	1	0

GATr leverages *projective geometric algebra* (PGA) to represent its features as 16-dimensional ***multivectors*** *encoding geometric objects*:

$$x = (\underbrace{x_s}_{\text{scalar}}, \underbrace{x_0, x_1, x_2, x_3}_{\text{vectors}}, \underbrace{x_{01}, x_{02}, x_{03}, x_{12}, x_{13}, x_{23}}_{\text{bi-vectors}}, \underbrace{x_{012}, x_{013}, x_{023}, x_{123}}_{\text{tri-vectors}}, \underbrace{x_{0123}}_{\text{pseudo-scalar}}) \in G(3,0,1)$$

Virtual Nodes Embeddings GATr

PerceiverIO learns fixed query tokens which attend to input tokens

Naive adaptation can *only capture invariant quantities*

$$\langle \boldsymbol{q}(v_i), \boldsymbol{k}(g.n_j)\rangle = \langle \boldsymbol{q}(v_i), \rho(g)\boldsymbol{k}(n_j)\rangle \quad \neq \langle \boldsymbol{q}(v_i), \boldsymbol{k}(n_j)\rangle \quad \forall g \in G$$

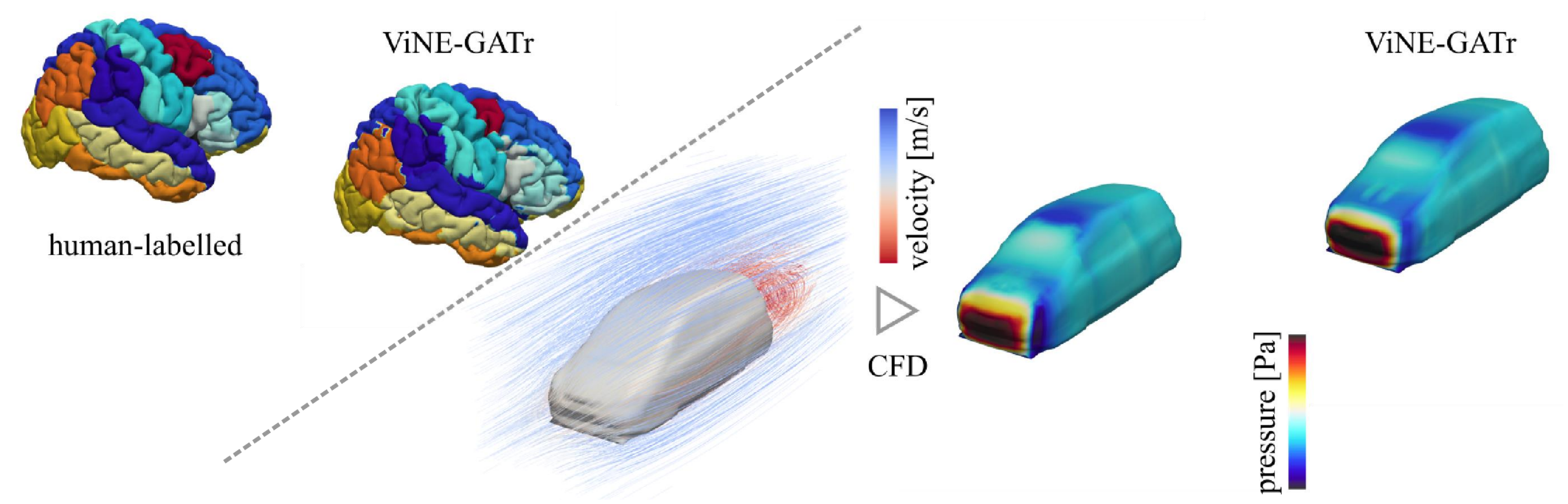
Virtual Nodes Embeddings **(ViNE)**:

- Augment query tokens with geometric features from input PCA
- Eigenvectors sign-ambiguity turned in permutation symmetry of right-handed frames by repeating virtual nodes 4 times
- It can learn a *canonicalized* virtual point-cloud by mixing features with GATr primitives

$$\langle \boldsymbol{q}(v_i, g.f), \boldsymbol{k}(g.n_j)\rangle = \langle \rho(g)\boldsymbol{q}(v_i, f), \rho(g)\boldsymbol{k}(n_j)\rangle \quad = \langle \boldsymbol{q}(v_i, f), \boldsymbol{k}(n_j)\rangle \quad \forall g \in G$$

$$X \in \mathbb{R}^{3\times N} \xrightarrow{\text{PCA}} \begin{matrix} \boldsymbol{c} \in \mathbb{R}^3 \\ [w_1 | w_2 | w_3] \in \mathbb{R}^{3\times 3} \end{matrix} \xrightarrow[\text{sign-ambiguity}]{\text{Resolve}} \begin{matrix} [+w_1 | +w_2 | +w_3] \\ [+w_1 | -w_2 | -w_3] \\ [-w_1 | -w_2 | +w_3] \\ [-w_1 | +w_2 | -w_3] \end{matrix}$$

Experiments



Large-scale cortical surface data

Compare ViNE-GATr (V=750 virtual nodes) with *furthest point sampling* (FPS) LaB-GATr and *random sampling* LaB−GATr† (~1K hidden tokens)

ViNE-GATr breaks eigenvector sign-flip symmetry

Model	Inference [ms] (∗)	Accuracy ↑ [%]
LaB-GATr	2076.5 (86.2 %)	**79.4** ± 1.8
ViNE-GATr	390.1 (1.8 %)	77.3 ± 2.7
LaB-GATr†	**386.1** (0.0 %)	66.6 ± 2.2

Mindboggle-101 dataset (~300K vertices per sample)

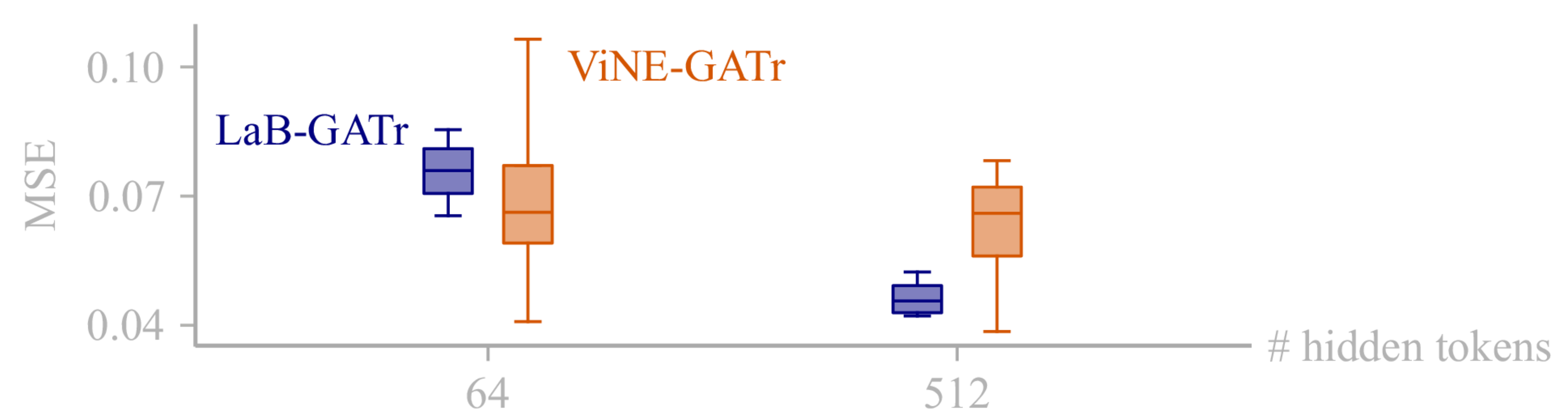
Cars' airflow data

Compare with LaB-GATr (FPS) when *varying the virtual token budget*

ViNE-GATr respects eigenvector sign-flip symmetry

Model	Error ×1e2 ↓
GINO	**2.14**
UPT	2.24
FNO	3.26
ViNE-GATr	3.85

ShapeNet Car benchmark



Bibliography

Brehmer et al. "Geometric Algebra Transformer" in NeurIPS, 2023.
Alkin et al. "Universal physics transformers: A framework for efficiently scaling neural operators" in NeurIPS, 2024.
Suk et al., "LaB-GATr: Geometric algebra transformers for large biomedical surface and volume meshes" in MICCAI 2024