TetraSphere

A Neural Descriptor for O(3)-Invariant Point Cloud Analysis

CVPR SEATTLE, WA JUNE 17-21, 2024

Pavlo Melnyk, Andreas Robinson, Michael Felsberg, Mårten Wadenbäck

Computer Vision Laboratory, Linköping University, Sweden

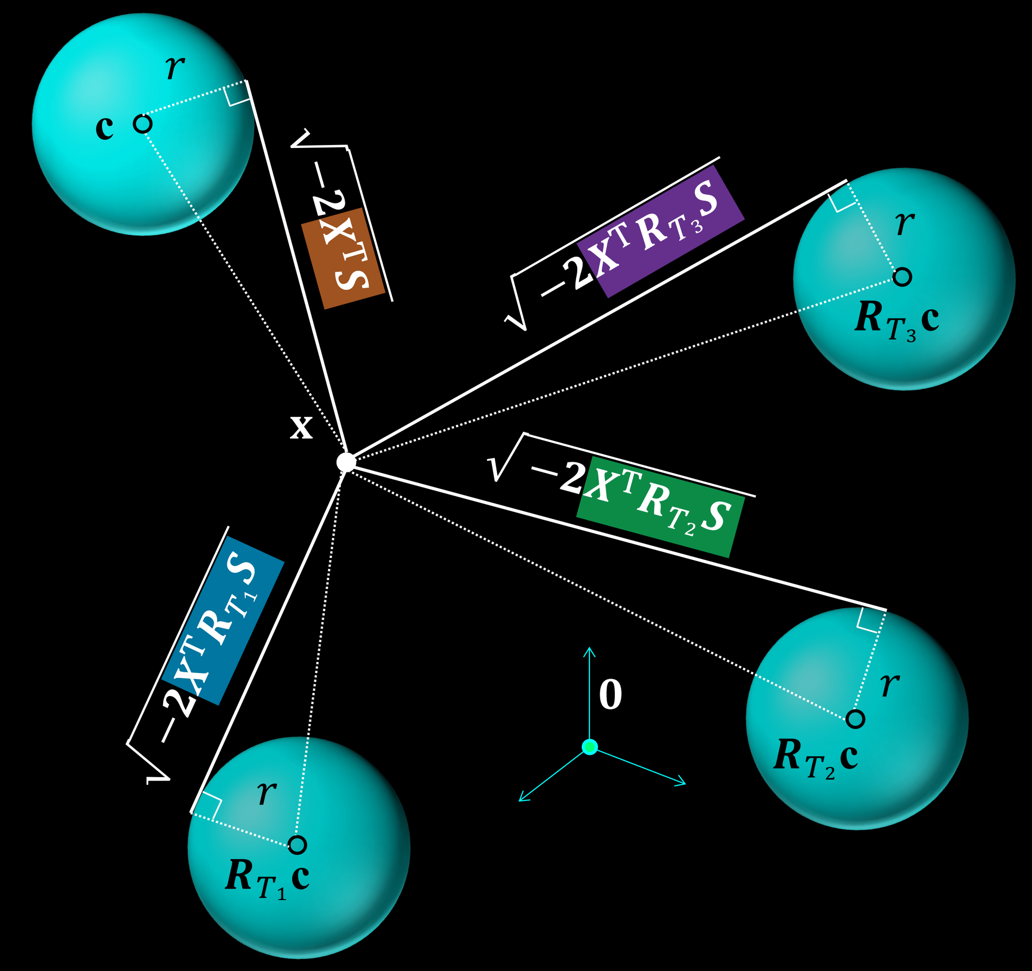
{ pavlo.melnyk, andreas.robinson, michael.felsberg, marten.wadenback }@liu.se

Motivation

- **Geometric** learning on 3D point clouds with NNs
- Permutation and O(3)-**equivariance** $F(\boldsymbol{R}\boldsymbol{X}) = \rho(\boldsymbol{R})\, F(\boldsymbol{X})$
- Requirement for natural science applications
- **Invariant** predictions if desired $F(\boldsymbol{R}\boldsymbol{X}) = F(\boldsymbol{X})$
- Eliminates the need for data augmentation

Background

- Two sorts of O(n)-equivariant neurons
- Steerable 3D spherical neurons [Melnyk et al., ICML'22] never used for end-to-end learning previously

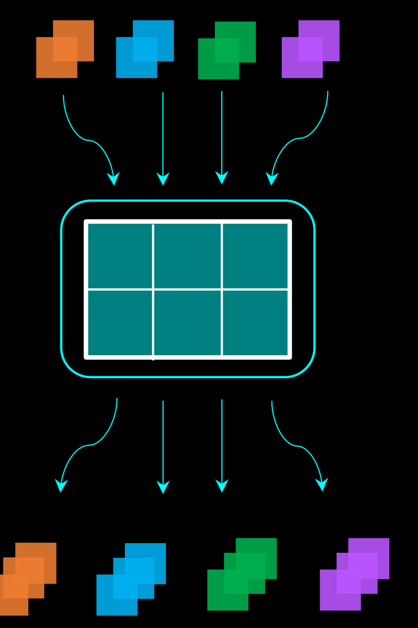


$$\boldsymbol{X} = (\mathrm{x}, -1, -\tfrac{1}{2}\|\mathrm{x}\|^2) \quad \in \mathbb{R}^5$$

$$\boldsymbol{S} = (\mathrm{c}, \tfrac{1}{2}(\|\mathrm{c}\|^2 - r^2), 1) \quad \in \mathbb{R}^5$$

$$\boldsymbol{RS} = (\boldsymbol{R}\mathrm{c}, \tfrac{1}{2}(\|\mathrm{c}\|^2 - r^2), 1) \quad \in \mathbb{R}^5$$

$$B(\boldsymbol{S})X = [\,\blacksquare\ \blacksquare\ \blacksquare\ \blacksquare\,]^{\mathrm{T}} \quad \in \mathbb{R}^4$$



$\boldsymbol{Y} \in \mathbb{R}^{C\times 4}$

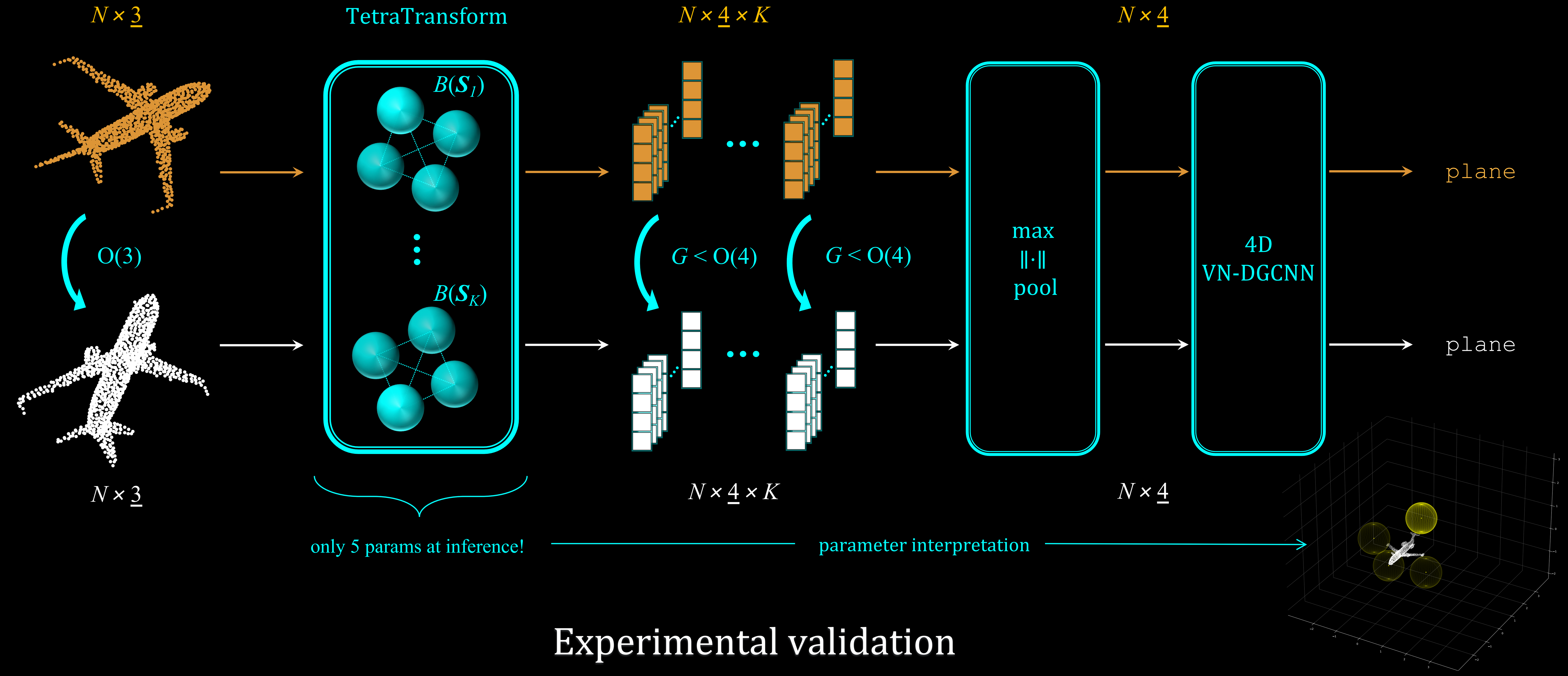
$\mathbf{W} \in \mathbb{R}^{C'\times C}$

$\boldsymbol{Y}' \in \mathbb{R}^{C'\times 4}$

- Vector neurons [Deng et al., ICCV'21]

Method

- TetraSphere := steerable 3D spherical neurons + equivariant pooling + 4D VN-DGCNN



Experimental validation

- SOTA classification results
- Randomly rotated **ScanObjectNN** real-world scans
- **Δ=0.7-1.3%** better than the previous best
- Decent on synthetic **ModelNet40** and **ShapeNet** (segmentation)
- TetraSphere > 4D VN-DGCNN > 3D VN-DGCNN

LINKÖPING UNIVERSITY

WASP WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM