Nonlinear Equivariant Neural Networks on Homogeneous Spaces

Elias Nyholm*, Oscar Carlsson, Daniel Persson
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
***eliasny@chalmers.se**

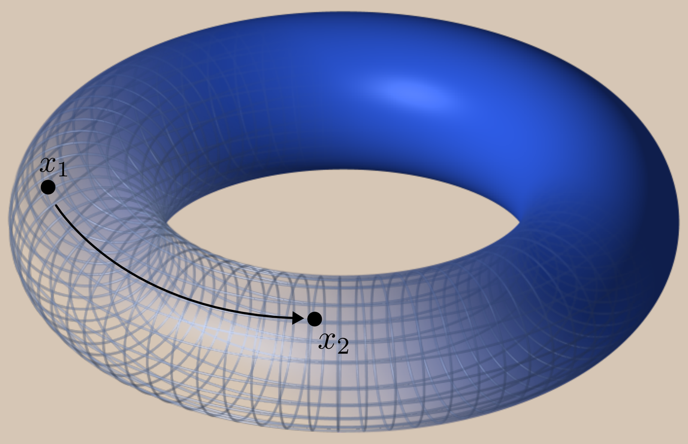
Summary

Our frameworks extends current families of equivariant **linear** neural network layers (G-CNNs, steerable CNNs) in order to include also **nonlinear** layers such as self-attention and message passing.

Background

Homogeneous Spaces

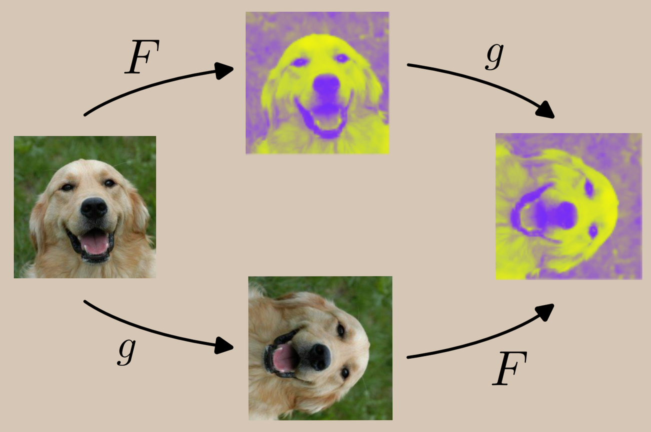
For all pairs of points $x_1, x_2 \in X$ there exist a transformation $g \in G$ such that $g \cdot x_1 = x_2$



Examples: Torus $\mathbb{T}^n$ (above), n-Sphere $\mathbb{S}^n$, Euclidean Space $\mathbb{E}^n$, Projective Space $\mathbb{R}P^n$

G-Equivariance

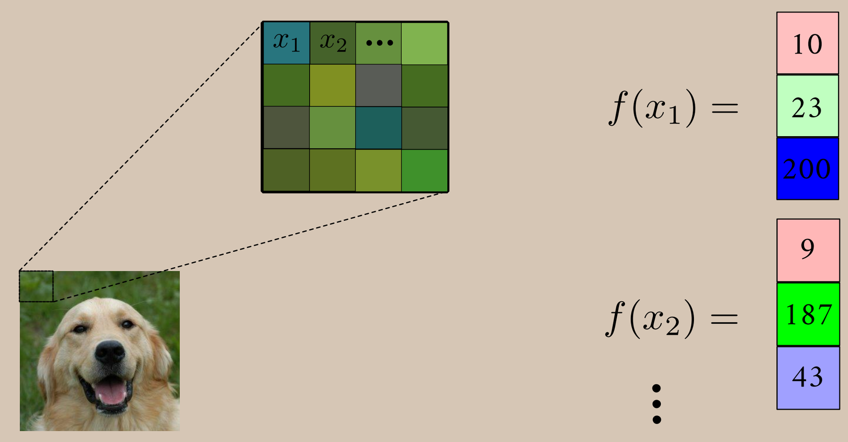
A map $F : X \to Y$ which commutes with all transformations $g \in G$ on X and Y



Examples: Inverting colors under rotations, CNN layers under translations, GNN networks under node permutations

Feature Maps

Data is encoded as a function $f : X \to \mathbb{R}^n$



Examples: Image ($X = \mathbb{E}^2, n = 3$), molecular data ($X = \mathbb{E}^3, n = 1$), graph with vector features ($X =$ graph)

Our proposal

We model neural network layers as operators between feature maps on homogeneous spaces, and study a family of G-equivariant operators on the form

$$\underbrace{f(g)}_{\text{vector}} \mapsto \int_G \underbrace{\omega(f, g, g')}_{\text{matrix}} \underbrace{f(g')}_{\text{vector}} dg' \qquad (1)$$

where we require that $\omega(f, g, g')$ only depends on $g^{-1}g', f(g), f(g')$ and that $\omega(f, gh, g') = \rho(h^{-1})\omega(f, g, g')$ for all $h \in H$.

Since the integration kernel parametrizes all such maps, we are interested in studying the space of such functions ω satisfying the condition above.

Special Cases

The framework specialises to existing equivariant models if one restricts the kernel further:

Condition	Yields
ω has no dependence on f	G-CNNs and steerable CNNs [Cohen et. al. 2018]
ω takes values in identity matrices times a scalar	LieTransformer [Hutchinson et. al. 2020]
ω ranges over all distributions	Integral map ranges over all equivariant maps

Properties of space of kernels

Two kernels ω, ω' such that

$$\int_H \omega(f, g, g'h)\rho(h^{-1})dh = \int_H \omega'(f, g, g'h)\rho(h^{-1})dh$$

for all f, g, g' yield the same integral operator in (1). This specialises to a known property of G-CNNs and an unknown property of the LieTransformer.

The structure of kernels is qualitatively similar to the structure of feature maps. In particular, many structures can be described as fibre bundles which yield geometric and topological information.

The space depends on the choice of reference function, e.g. Legesgue integrable functions L^1 or distributions (generalised functions). The standard choice in the literature is the space of square-integrable functions L^2.

Conclusions

Our framework generalizes established equivariant models. Any result proved for ansatz (1) yields corresponding results for each specialisation.

The framework is powerful in the sense that it uncovers mathematical structures in general equivariant operators, such as fibre bundles, which yield constraints on e.g. the expressivety of equivariant models.

Further studying this family of operators might give insights into the design of next-generation equivariant machine learning models, including equivariant transformer models.

WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM
CHALMERS UNIVERSITY OF TECHNOLOGY | UNIVERSITY OF GOTHENBURG