A (MORE) GENERAL FRAMEWORK FOR EQUIVARIANT NEURAL NETWORKS

ELIAS NYHOLM
DEPARTMENT OF MATHEMATICAL SCIENCES
CHALMERS UNIVERSITY OF TECHNOLOGY & UNIVERSITY OF GOTHENBURG

WORK WITH OSCAR CARLSSON, MAURICE WEILER AND DANIEL PERSSON. IN PREPARARTION: ARXIV:2503.*****

ABSTRACT

GOAL: WE WANT TO (MATHEMATICALLY) UNIFY ML MODELS IN ORDER TO SHOW SIMILARITIES AND DIFFERENCES BETWEEN THEM, AND POTENTIALLY CHARACTERISE THESE FEATURES USING SYMMETRY ARGUMENTS

EXAMPLE:
CONVOLUTIONAL LAYERS = LINEARITY + TRANSLATIONAL SYMMETRY

OUTLINE:

BACKGROUND → OUR FRAMEWORK → CONCLUSIONS

REPRESENTING DATA AND MODEL

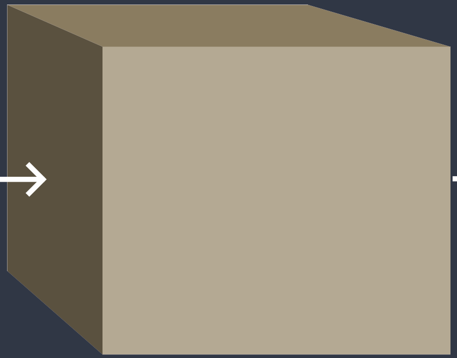
TRANSFORMING
DATA

$$\mathcal{F} \ni f : X \to \mathbb{R}^n$$

W/ GROUP ACTION

$$(g, f) \mapsto g \cdot f$$

EQUIVARIANT
MODEL



$$\Phi_\omega : \mathcal{F} \to \mathcal{F}$$

SATISFYING EQUIVARIANCE

$$\Phi[g \cdot f] = g \cdot \Phi[f]$$

TRANSFORMING
DATA

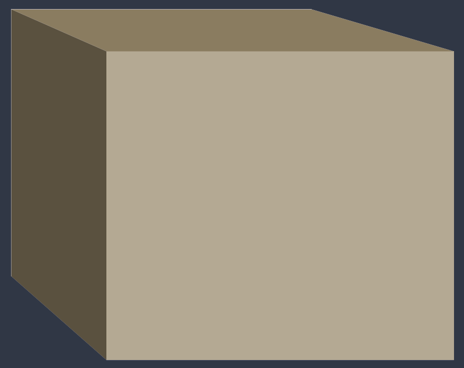
$$\mathcal{F} \ni f : X \to \mathbb{R}^n$$

W/ GROUP ACTION

$$(g, f) \mapsto g \cdot f$$

EQUIVARIANT ARCHITECTURES

EQUIVARIANT
MODEL



CONVOLUTIONAL NEURAL NETWORKS
G-CNNS AND STEERABLE CNNS

TRANSFORMER NETWORKS
LIETRANSFORMER

OTHER ARCHITECTURES...
GRAPH NEURAL NETWORKS, ...

ANSATZ' FOR OTHER EQUIVARIANT MODELS

G-CNN

$$\Phi_\kappa[f](g) = \int_G \kappa(g, g') f(g') dg'$$

$$\kappa(g, g') \in \mathbb{R}^{n \times m}$$

EQUIVARIANT IF

$$\kappa(g, g') = \kappa(g^{-1} g')$$

LIETRANSFORMER

$$\Psi_\omega[f](g) = \int_G \omega_f(g, g') f(g') dg'$$

$$\omega_f(g, g') \in \mathbb{R}$$

EQUIVARIANT IF

$$\omega_f(g, g') = \frac{\langle W^Q f(g), W^K f(g') \rangle}{\sqrt{N}} + \phi(g^{-1} g')$$

[1] Cohen, Taco, and Welling, Max. "Group Equivariant Convolutional Neural Networks", International Conference on Machine Learning, 2016.
[2] Hutchinson, Michael J., Charline Le Lan, Sheheryar Zaidi, Emilien Dupont, Yee Whye Teh and Hyunjik Kim. “LieTransformer: Equivariant self-attention for Lie Groups.” International Conference on Machine Learning, 2020.

OUR PROPOSAL

$$\underbrace{\Phi_\omega[f](g)}_{\text{vector}} = \int_G \underbrace{\omega(f, g, g')}_{\text{matrix}} \underbrace{f(g')}_{\text{vector}} dg'$$

EQUIVARIANCE CONDITION

$$\Phi_\omega[f](g) = \int_G \omega(f, g, g') f(g') dg'$$

$\downarrow$ **EQUIVARIANT IF**

$$\omega(f, g, g') \text{ is a function of } f(g), f(g'), g^{-1}g' \ldots$$

CONCLUSION 1: A TRANSFORMER-LIKE DEPENDENCE ARISES NATURALLY FROM EQUIVARIANCE, WITH RELATIVE POSITIONAL ENCODING ON THE TRANSFORMER SIDE DIRECTLY CORRESPONDING TO THE KERNEL OF CONVOLUTIONAL NETWORKS.

RELATION TO OTHER ARCHITECTURES

$$\Phi_\omega[f](g) = \int_G \omega(f, g, g')f(g')dg'$$

$\omega(f, g, g') \in \mathbb{R}$

LIETRANSFORMER

$\omega(f, g, g') = \kappa(g, g')$

G-CNN

$\omega(f, g, g') = \delta(g^{-1}g')\lambda[f](g)\frac{f(g')^T}{\|f(g')\|^2}$

ARBITRARY EQUIVARIANT OPERATOR

CONCLUSION 2: THE PROPOSED FRAMEWORK IS FLEXIBLE, GENERALISES NOT ONLY WELL-KNOWN PRE-EXISTING EQUIVARIANT ARCHITECTURES BUT CAN IN FACT DESCRIBE ANY EQUIVARIANT LAYER/OPERATOR.

MATHEMATICAL STRUCTURES

$$\omega(f, g, g')$$

- FORM A SPACE OF SECTIONS OF SOME EQUIVARIANT BUNDLE

- IS AN INDUCED REPRESENTATION OF SOME SUBRGOUP H OF G

- IS ASSOCIATED TO AN EQUIVALENCE CLASS OF MAPS WHICH ALL RESULT IN THE SAME OPERATOR

CONCLUSION 3: WE HAVE SEVERAL WAYS OF CHARACTERISING THE SPACE OF OPERATORS. ANALYSIS THESE SPACES FURTHER MIGHT GIVE US INSIGHTS IN HOW LARGE THIS SPACE IS, ITS TOPOLOGY ETC.

SUMMARY

CONCLUSION 1: A TRANSFORMER-LIKE DEPENDENCE ARISES NATURALLY FROM EQUIVARIANCE, WITH RELATIVE POSITIONAL ENCODING ON THE TRANSFORMER SIDE DIRECTLY CORRESPONDING TO THE KERNEL OF CONVOLUTIONAL NETWORKS.

CONCLUSION 2: THE PROPOSED FRAMEWORK IS FLEXIBLE, GENERALISES NOT ONLY WELL-KNOWN PRE-EXISTING EQUIVARIANT ARCHITECTURES BUT CAN IN FACT DESCRIBE ANY EQUIVARIANT LAYER/OPERATOR.

CONCLUSION 3: WE HAVE SEVERAL WAYS OF CHARACTERISING THE SPACE OF OPERATORS. ANALYSIS THESE SPACES FURTHER MIGHT GIVE US INSIGHTS IN HOW LARGE THIS SPACE IS, ITS TOPOLOGY ETC.

THANK YOU!

QUESTIONS?