Geometric deep learning on the sphere

Efficient generalised spherical CNNs

Jason McEwen

www.jasonmcewen.org

Kagenova Limited

Mullard Space Science Laboratory (MSSL), UCL

In collaboration with:

Oliver Cobb · Chris Wallis · Augustine Mavor-Parker · Augustin Marignier · Matthew Price · Mayeul d'Avezac

20 April 2021

Mentimeter

Give your input at `https://www.menti.com/puiqjn97i9`.

Or go to `https://www.menti.com` and enter voting code: 80 20 22 0.

Outline

1. Symmetry in deep learning
2. Spherical CNNs
3. Efficient generalised spherical CNNs
4. Numerical results

Symmetry in deep learning

Physics and deep learning

Physics

Understanding the world by
modelling from first principles
for generative models and inference.

Deep Learning

Understanding the world by
learning informative representations
for generative models and inference.

Physics and deep learning

Physics

Understanding the world by
modelling from first principles
for generative models and inference.

Hard!

Deep Learning

Understanding the world by
learning informative representations
for generative models and inference.

Hard!

Physics and deep learning

Physics $\longleftrightarrow$ Deep Learning

Physics and deep learning

Physics $\longleftrightarrow$ Deep Learning

Here we focus on integrating physics $\rightarrow$ deep learning

(in other works focus on reverse: physics $\leftarrow$ deep learning).

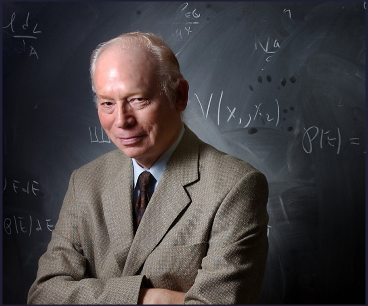
Physics and deep learning

Physics $\longleftrightarrow$ Deep Learning

Here we focus on integrating physics $\rightarrow$ deep learning

(in other works focus on reverse: physics $\leftarrow$ deep learning).

As we will see, this key factor driving the deep learning revolution.



“Symmetry: key to nature’s secrets.”

— Steven Weinberg

Symmetry

Mirror symmetry

Symmetry

Mirror symmetry

Symmetry

Mirror symmetry

Rotational symmetry

Symmetry

Mirror symmetry

Rotational symmetry

Symmetry

Mirror symmetry

Rotational symmetry

Symmetry

Mirror symmetry

Rotational symmetry

Symmetry

Mirror symmetry

Rotational symmetry

Symmetry (invariance) to continuous transformation

In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.

Spatial translation

Symmetry (invariance) to continuous transformation

In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.

Spatial translation

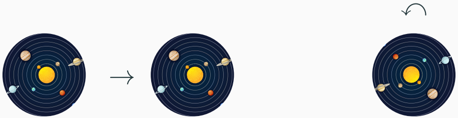
In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.

Spatial translation

Rotation

Symmetry (invariance) to continuous transformation

In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.

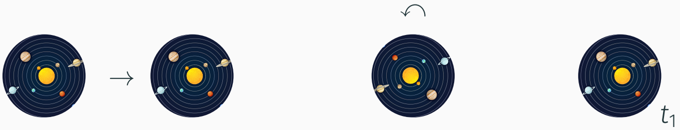


Spatial translation

Rotation

Symmetry (invariance) to continuous transformation

In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.

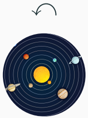


Spatial translation Rotation Time translation

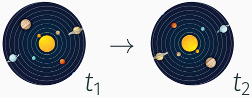
Symmetry (invariance) to continuous transformation

In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.

Spatial translation



Rotation



Time translation

Mentimeter

Give your input at `https://www.menti.com/puiqjn97i9`.

Or go to `https://www.menti.com` and enter voting code: 80 20 22 0.

Noether's theorem

Noether's theorem

For every continuous symmetry of the universe, there exists a conserved quantity.

Emmy Noether

Noether's theorem

Noether's theorem

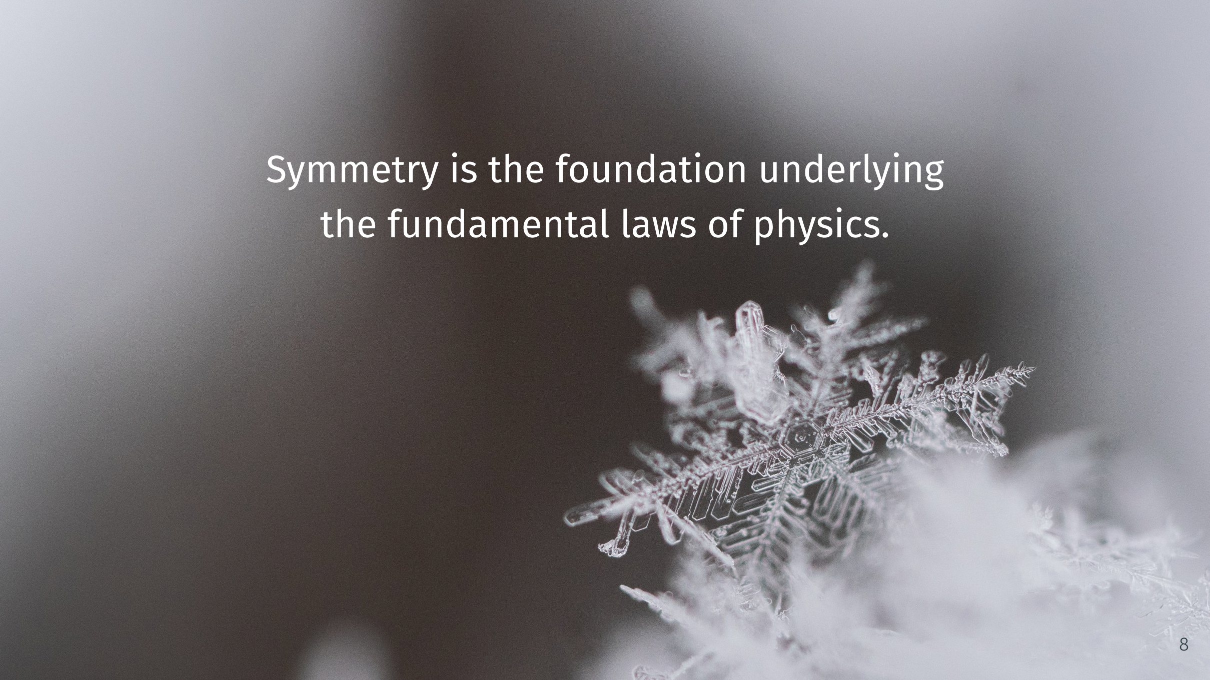
For every continuous symmetry of the universe, there exists a conserved quantity.

Symmetries at the heart of physics:

- **Translational** symmetry ⇔ conservation of **momentum**
- **Rotational** symmetry ⇔ conservation of **angular momentum**
- **Time translational** symmetry ⇔ conservation of **energy**

(Energy not conserved in general relativity since time translation broken.)

Emmy Noether


Symmetry is the foundation underlying
the fundamental laws of physics.

Symmetry in deep learning

Encoding symmetry in deep learning models captures fundamental properties about the underlying nature of our world.

Symmetry in deep learning

Encoding symmetry in deep learning models captures fundamental properties about the underlying nature of our world.

Key factor driving the deep learning revolution, with the advent of CNNs.

- CNNs resulted in a step-change in performance.
- Convolutional structure of CNNs capture translational symmetry (i.e. translational equivariance).

Mentimeter

Give your input at `https://www.menti.com/puiqjn97i9`.

Or go to `https://www.menti.com` and enter voting code: 80 20 22 0.

Equivariance

Equivariance

An operator $\mathcal{A}$ is equivariant to a transformation $\mathcal{T}$ if

$$\mathcal{T}\big(\mathcal{A}(f)\big) = \mathcal{A}\big(\mathcal{T}(f)\big)$$

for all possible signals f.

Transforming the signal after application of the operator, is equivalent to transformation of the signal first, followed by application of the operator.

Equivariance

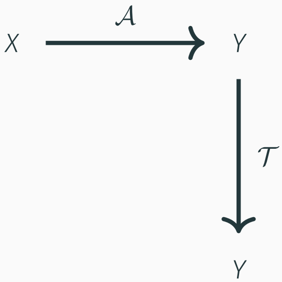
Equivariance

An operator $\mathcal{A}$ is equivariant to a transformation $\mathcal{T}$ if

$$\mathcal{T}\big(\mathcal{A}(f)\big) = \mathcal{A}\big(\mathcal{T}(f)\big)$$

for all possible signals f.

Transforming the signal after application of the operator, is equivalent to transformation of the signal first, followed by application of the operator.



Equivariance

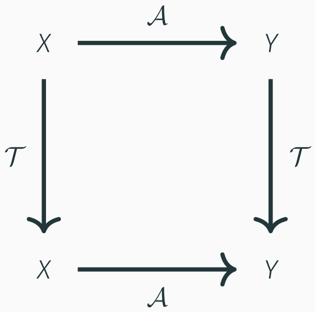
Equivariance

An operator $\mathcal{A}$ is equivariant to a transformation $\mathcal{T}$ if

$$\mathcal{T}\big(\mathcal{A}(f)\big) = \mathcal{A}\big(\mathcal{T}(f)\big)$$

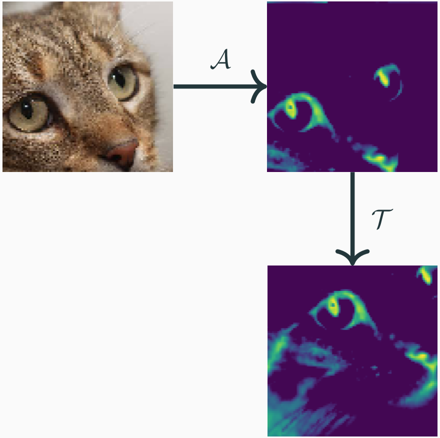
for all possible signals f.

Transforming the signal after application of the operator, is equivalent to transformation of the signal first, followed by application of the operator.



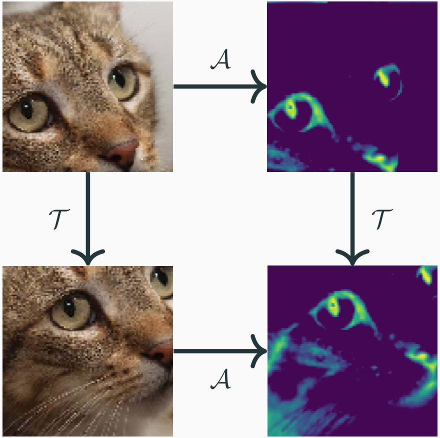
Planar (Euclidean) CNNs exhibit translational equivariance

Planar (Euclidean) convolution is translationally equivariant.



Planar (Euclidean) CNNs exhibit translational equivariance

Planar (Euclidean) convolution is translationally equivariant.



Importance of equivariance

Imposing inductive biases in deep learning models, such as equivariance to symmetry transformations, allows models to be learned in a more principled and effective manner.

Capture fundamental physical understanding of generative process.

Importance of equivariance

In some sense, equivariance to a transformation means a pattern need only be learnt once, and may then be recognised in all transformed scenarios.

Importance of equivariance

In some sense, equivariance to a transformation means a pattern need only be learnt once, and may then be recognised in all transformed scenarios.

Cat

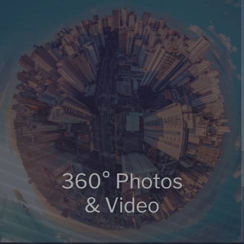
Still a cat

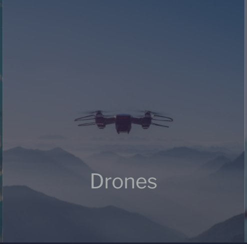
Spherical CNNs

Data on the sphere is prevalent

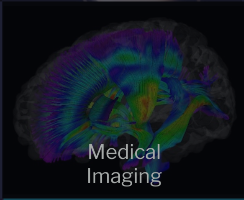
Data on the sphere is prevalent

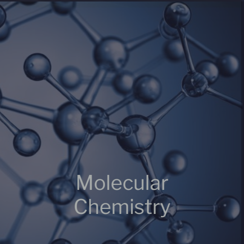
Encode symmetries of the sphere and rotations

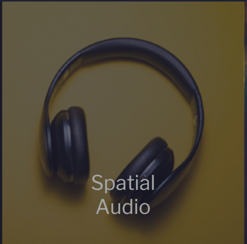


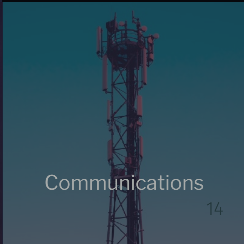


Data on the sphere arises in many applications



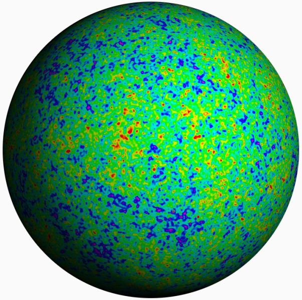






Cosmology and virtual reality

Whenever observe over angles, recover data on 2D sphere (or 3D rotation group).

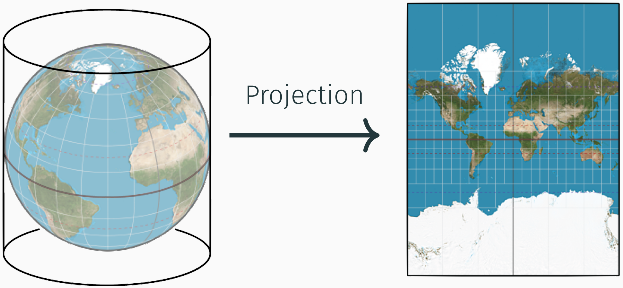


Cosmic microwave background

360° virtual reality

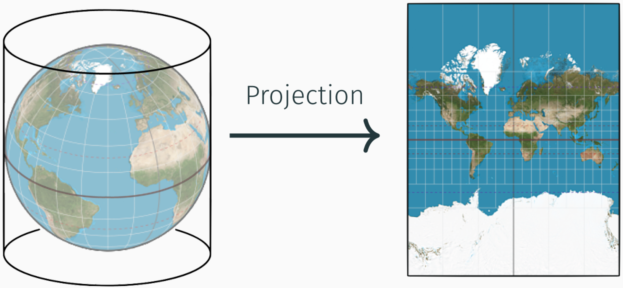
Why not standard (Euclidean) deep learning approaches?

Could project sphere to plane and then apply standard planar CNNs.



Why not standard (Euclidean) deep learning approaches?

Could project sphere to plane and then apply standard planar CNNs.



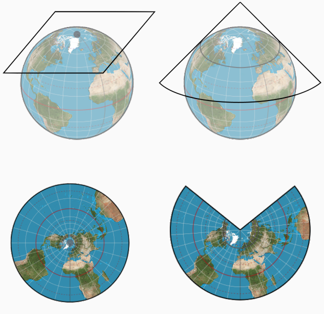
Greenland appears to be a similar size to Africa in the projected planar map, whereas it is over 10 times smaller.

Why not standard (Euclidean) deep learning approaches?

Projection breaks symmetries and geometric properties of sphere.

Why not standard (Euclidean) deep learning approaches?

Projection breaks symmetries and geometric properties of sphere.

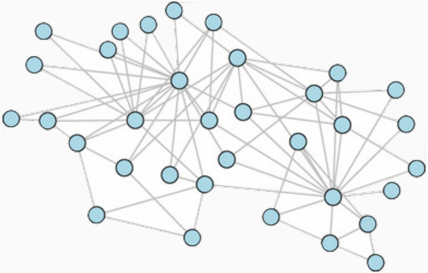


No projection of the sphere to the plane can preserve both shapes and areas
$\Rightarrow$ distortions are unavoidable.

(Formally: a conformal, area-preserving projection does not exist.)

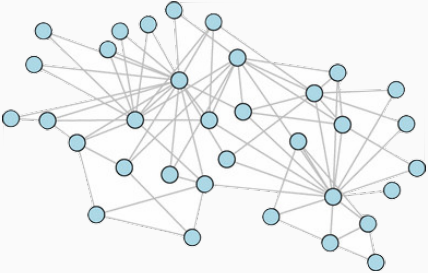
Why not graph-based geometric deep learning?

Could construct graph representation of sphere and apply graph CNNs.



Why not graph-based geometric deep learning?

Could construct graph representation of sphere and apply graph CNNs.

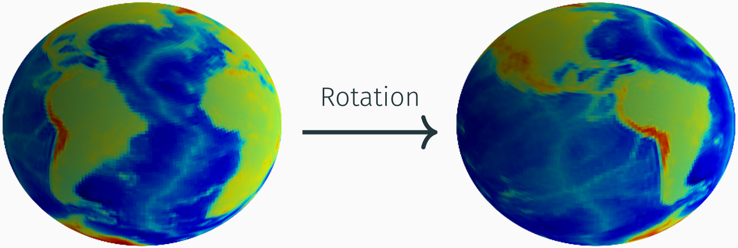


Again, **breaks symmetries and geometric properties** of sphere.

Cannot capture rotational equivariance.

Rotational equivariance

On the sphere, the analog of translations are rotations.



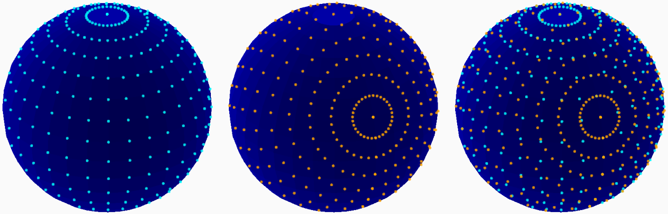
Would like spherical CNNs to exhibit rotational equivariance.

(Just as planar CNNs exhibit translational equivariance.)

Capturing rotational equivariance in spherical CNNs

Well-known that regular discretisation of the sphere does not exist (e.g. Tegmark 1996).

$\Rightarrow$ Not possible to discretise sphere in a manner that is invariant to rotations.

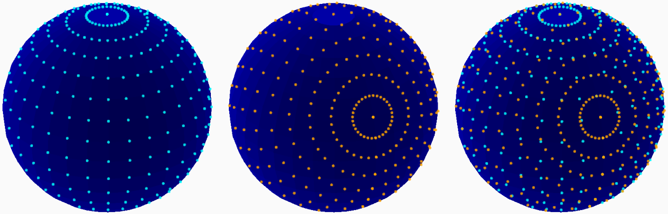


Capturing strict equivariance with operations defined directly in discretised (pixel) space **not possible** due to structure of sphere.

Capturing rotational equivariance in spherical CNNs

Well-known that regular discretisation of the sphere does not exist (e.g. Tegmark 1996).

$\Rightarrow$ Not possible to discretise sphere in a manner that is invariant to rotations.

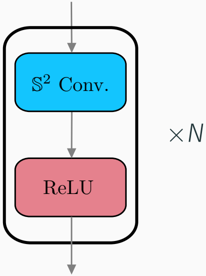


Capturing strict equivariance with operations defined directly in discretised (pixel) space **not possible** due to structure of sphere.

Instead, consider Fourier approach $\rightarrow$ access to underlying continuous representations.

Spherical CNN

Spherical CNNs constructed by analog of Euclidean CNNs but using convolution on the sphere and with pointwise non-linear activations functions, e.g. ReLU (Cohen et al. 2018; Esteves et al. 2018).



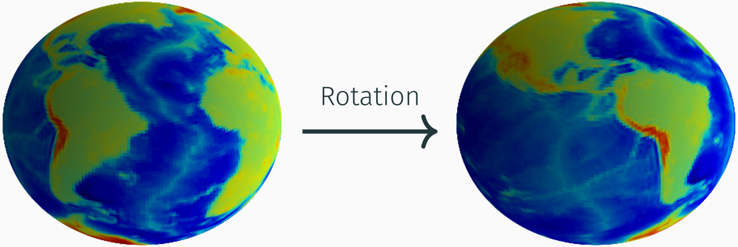
(Alternative, real space constructions have also been developed but do not exhibit rotational equivariance so not considered further; e.g. Boomsma & Frellsen 2017, Jiang et al. 2019, Perraudin et al. 2019.)

Rotation of signals

Rotation of signals in spatial domain

A signal $f \in L^2(\Omega)$ on the sphere ($\Omega = \mathbb{S}^2$) or rotation group ($\Omega = \mathrm{SO}(3)$) can be rotated by $\rho \in \mathrm{SO}(3)$ by

$$\mathcal{R}_\rho f(\omega) = f(\rho^{-1}\omega), \qquad \text{for } \omega \in \Omega.$$



Convolution of signals

Convolution of signals in spatial domain

Convolution of two signals $f, \psi \in L^2(\Omega)$ is given by

$$(f \star \psi)(\rho) = \langle f, \mathcal{R}_\rho \psi \rangle = \int_\Omega \mathrm{d}\mu(\omega)\, f(\omega)\, \psi^*(\rho^{-1}\omega), \quad \text{for } \omega \in \Omega, \rho \in \mathrm{SO}(3),$$

where $d\mu(\omega)$ denotes the Haar measure on Ω and $\cdot^*$ complex conjugation.

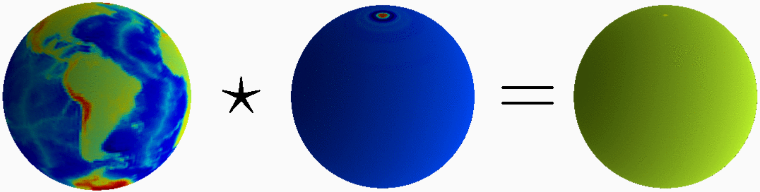
Convolution of signals

Convolution of signals in spatial domain

Convolution of two signals $f, \psi \in L^2(\Omega)$ is given by

$$(f \star \psi)(\rho) = \langle f, \mathcal{R}_\rho \psi \rangle = \int_\Omega \mathrm{d}\mu(\omega)\, f(\omega)\, \psi^*(\rho^{-1}\omega), \quad \text{for } \omega \in \Omega, \rho \in \mathrm{SO}(3),$$

where $d\mu(\omega)$ denotes the Haar measure on Ω and $\cdot^*$ complex conjugation.



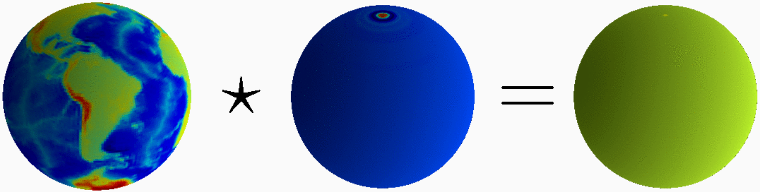
Convolution of signals

Convolution of signals in spatial domain

Convolution of two signals $f, \psi \in L^2(\Omega)$ is given by

$$(f \star \psi)(\rho) = \langle f, \mathcal{R}_\rho \psi \rangle = \int_\Omega \mathrm{d}\mu(\omega)\, f(\omega)\, \psi^*(\rho^{-1}\omega), \qquad \text{for } \omega \in \Omega, \rho \in \mathrm{SO}(3),$$

where $d\mu(\omega)$ denotes the Haar measure on Ω and $\cdot^*$ complex conjugation.

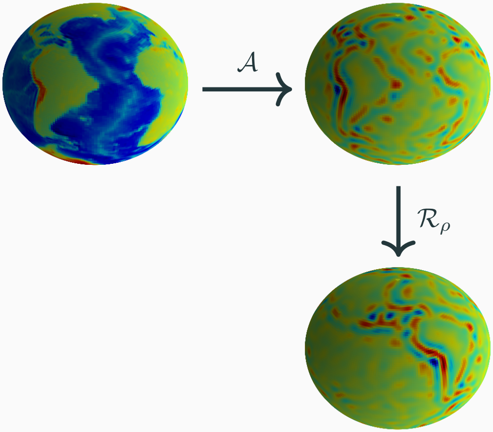


Since no regular discretization of the sphere, compute in Fourier space to ensure equivariant.

Convolution is rotationally equivariant

Convolution is rotational equivariant (when computed in harmonic domain):

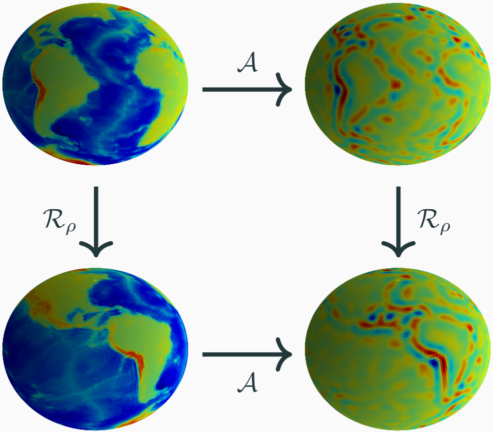
$$((\mathcal{R}_\rho f) \star \psi)(\rho') = (\mathcal{R}_\rho (f \star \psi))(\rho').$$



Convolution is rotationally equivariant

Convolution is rotational equivariant (when computed in harmonic domain):

$$((\mathcal{R}_\rho f) \star \psi)(\rho') = (\mathcal{R}_\rho (f \star \psi))(\rho').$$



Pointwise activation

While pointwise activations are rotationally equivariant in the continuous limit, they are not equivariant in practice when applied to discretised signals (since regular discretisation of sphere does not exist).

Pointwise activation

While pointwise activations are rotationally equivariant in the continuous limit, they are not equivariant in practice when applied to discretised signals (since regular discretisation of sphere does not exist).

Equivariance errors

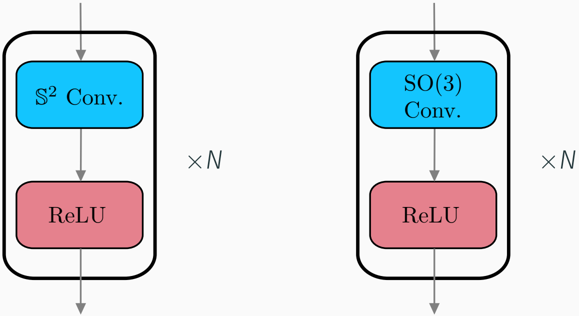
Layer	Mean Relative Error*
$\mathbb{S}^2$ to $\mathbb{S}^2$ conv.	4.4×10^{-7}
$\mathbb{S}^2$ to SO(3) conv.	5.3×10^{-7}
SO(3) to SO(3) conv.	9.3×10^{-7}
$\mathbb{S}^2$ ReLU	3.4×10^{-1}
SO(3) ReLU	3.7×10^{-1}

* Floating point precision.

Spherical CNNs

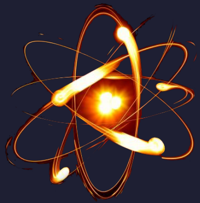
Approach taken by Cohen et al. 2018 and Esteves et al. 2018.

Despite imperfect equivariance, find empirically that such models maintain a reasonable degree of equivariance and generally perform well.



Efficient generalised spherical CNNs

Group theory is the mathematical study of symmetry.



Since we’re concerned with rotational symmetry, leverage the machinery from the study of angular momentum in quantum mechanics.

Generalized spherical CNNs

Consider the s-th layer of a generalized spherical CNN to take the form of a triple (Cobb et al. 2021)

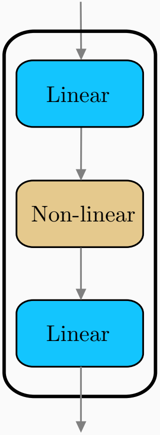
$$\mathcal{A}^{(s)} = (\mathcal{L}_1, \mathcal{N}, \mathcal{L}_2),$$

such that

$$\mathcal{A}^{(s)}(f^{(s-1)}) = \mathcal{L}_2(\mathcal{N}(\mathcal{L}_1(f^{(s-1)}))),$$

where

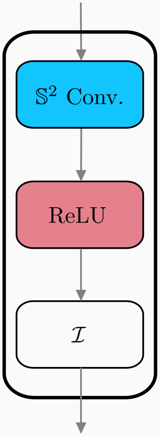
- $\mathcal{L}_1, \mathcal{L}_2 : \mathcal{F}_L \rightarrow \mathcal{F}_L$ are spherical convolution operators,
- $\mathcal{N} : \mathcal{F}_L \rightarrow \mathcal{F}_L$ is a non-linear, spherical activation operator.



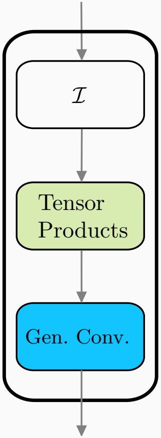
Generalised spherical CNNs

- Build on other **influential equivariant spherical CNN** constructions:
    - Cohen et al. (2018)
    - Esteves et al. (2018)
    - Kondor et al. (2018)
- Encompass other frameworks as special cases.
- General framework supports hybrids models.

Existing spherical CNN layers are **highly computationally costly**, particularly those non-linear layers that satisfy strict rotational equivariance.



Cohen et al. (2018), Esteves et al. (2018)



Kondor et al. (2018)

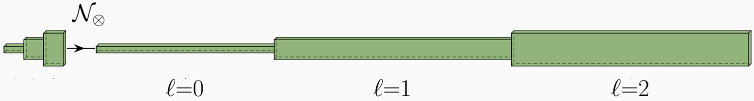
Contributions to improve efficiency

1. Channel-wise structure
2. Constrained generalized convolutions
3. Optimized degree mixing sets
4. Efficient sampling theory on the sphere and rotation group

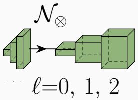
Channel-wise structure

Split generalized signals in K channels and apply a tensor-product activation to each channel separately.

Representational capacity then controlled through linear dependence on channels K, rather than quadratic dependence (on generalized harmonic representation type τ_f).



Prior approach to applying a tensor-product based non-linear operator



Ours (Cobb et al. 2021)

Constrained generalized convolutions

Under new multi-channel structure, decompose the generalized convolution into **three separate constrained linear operators**:

1. **Uniform convolution**: linear projection uniformly across channels to project down onto the desired type (interpreted as learned extension of tensor-product activations to undo expansion of representation space).
2. **Channel-wise convolution**: linear combinations of the fragments within each channel.
3. **Cross-channel convolution**: linear combinations to learn new features.

Computational and parameter efficiency significantly improved.

Optimized degree mixing sets

Non-linear operators must perform degree mixing (equivariant linear operators cannot mix information corresponding to different degrees).

But, it is **not necessary to compute all possible tensor-product based fragments**.

Degree mixing set $\mathbb{P}_L^\ell$:

$$\mathbb{P}_L^\ell = \{(\ell_1, \ell_2) \in \{0, ..., L-1\}^2 : |\ell_1 - \ell_2| \leq \ell \leq \ell_1 + \ell_2\}.$$

Consider subsets of $\mathbb{P}_L^\ell$ that scale better than $\mathcal{O}(L^2)$.

Optimized degree mixing sets

Consider the graph $G_L^\ell = (\mathbb{N}_L, \mathbb{P}_L^\ell)$ with nodes $\mathbb{N}_L = \{0, ..., L-1\}$ and edges $\mathbb{P}_L^\ell$.

- Some notion of relationship between ℓ_1 and ℓ_2 is captured if there exists a path between the two nodes in G_L^ℓ.
- Select smallest subgraph such that all relationships are preserved $\Rightarrow$ **minimum spanning tree** (MST). Weight edges by computational cost to minimise overall cost.
- Consider **logarithmic subsampling** (reduced MST).

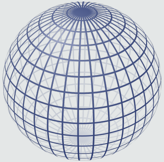
Computational complexity significantly reduced from $\mathcal{O}(L^5)$ to $\mathcal{O}(L^3 \log L)$, where L denotes resolution (bandlimit).

Efficient sampling theory and fast harmonic transforms

Adopt **efficient sampling theory** and **fast algorithms** to compute harmonic transforms on the sphere and rotation group.

Leverage to access underlying continuous signal representations, avoiding discretization artifacts, and compute fast convolutions.

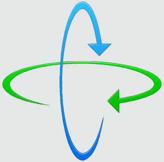
Novel sampling theorem on sphere (McEwen & Wiaux 2011)



SSHT: Spin spherical harmonic transforms

www.spinsht.org

Novel sampling theorem on rotation group (McEwen et al. 2015)

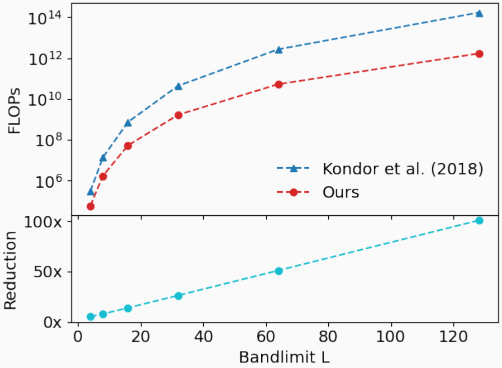


SO3: Fast Wigner transforms on rotation group

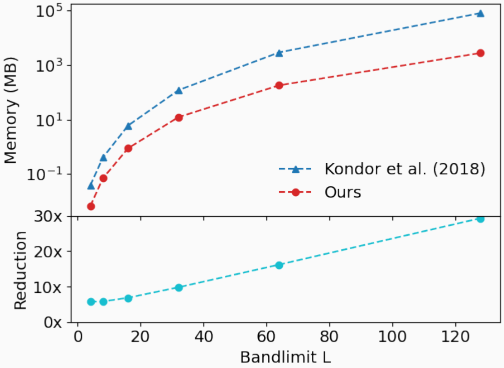
www.sothree.org

Numerical results

Computational cost and memory requirements



Computational cost



Memory requirements

Rotational equivariance

Equivariance errors

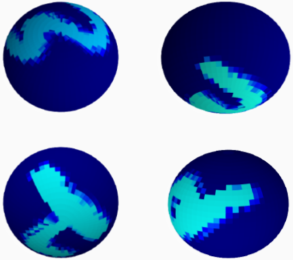
Layer	Mean Relative Error*
Tensor-product activation $\rightarrow$ Generalized convolution	5.0×10^{-7}
$\mathbb{S}^2$ ReLU	3.4×10^{-1}
SO(3) ReLU	3.7×10^{-1}

* Floating point precision.

Spherical MNIST: problem

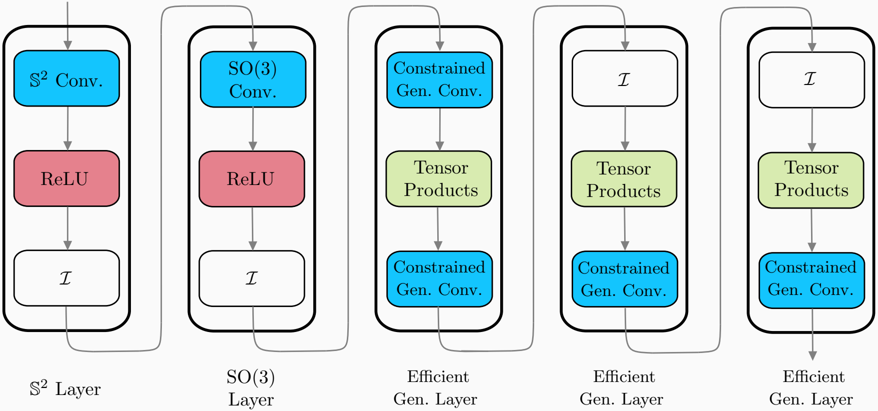
Canonical benchmark problem of classifying MNIST digits projects onto the sphere.

Non-rotated (NR)



Rotated (R)

Spherical MNIST: architecture



Spherical MNIST: results

Test accuracy for spherical MNIST digits classification problem

	NR/NR	R/R	NR/R	Params
Planar CNN	99.32			58k
Cohen et al. 2018	95.59			58k
Kondor et al. 2018	96.40			286k
Esteves et al. 2018	**99.37**			58k
Ours (MST)	99.35			58k
Ours (RMST)	99.29			**57k**

Spherical MNIST: results

Test accuracy for spherical MNIST digits classification problem

	NR/NR	R/R	NR/R	Params
Planar CNN	99.32	90.74		58k
Cohen et al. 2018	95.59	94.62		58k
Kondor et al. 2018	96.40	96.60		286k
Esteves et al. 2018	**99.37**	99.37		58k
Ours (MST)	99.35	**99.38**		58k
Ours (RMST)	99.29	99.17		**57k**

Spherical MNIST: results

Test accuracy for spherical MNIST digits classification problem

	NR/NR	R/R	NR/R	Params
Planar CNN	99.32	90.74	11.36	58k
Cohen et al. 2018	95.59	94.62		58k
Kondor et al. 2018	96.40	96.60		286k
Esteves et al. 2018	**99.37**	99.37		58k
Ours (MST)	99.35	**99.38**		58k
Ours (RMST)	99.29	99.17		**57k**

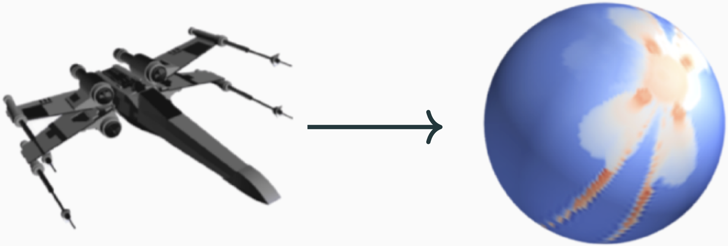
Spherical MNIST: results

Test accuracy for spherical MNIST digits classification problem

	NR/NR	R/R	NR/R	Params
Planar CNN	99.32	90.74	11.36	58k
Cohen et al. 2018	95.59	94.62	93.40	58k
Kondor et al. 2018	96.40	96.60	96.00	286k
Esteves et al. 2018	**99.37**	99.37	99.08	58k
Ours (MST)	99.35	**99.38**	**99.34**	58k
Ours (RMST)	99.29	99.17	99.18	**57k**

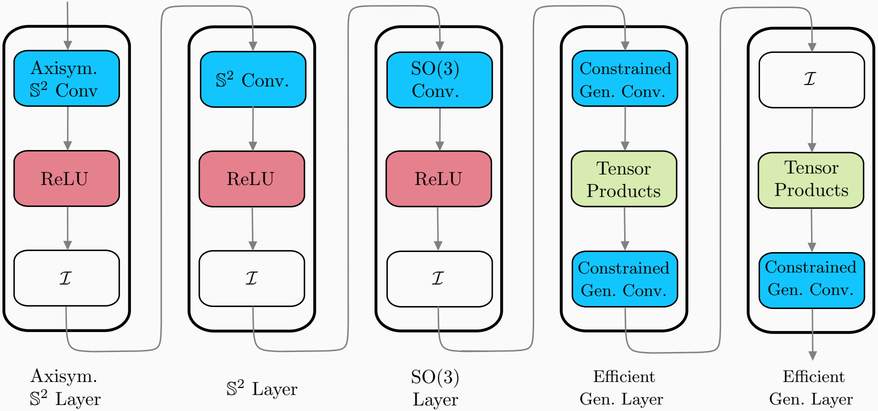
3D shape classification: problem

Classify 3D meshes and perform shape retrieval.



[Image credit: Esteves et al. 2018]

3D shape classification: architecture



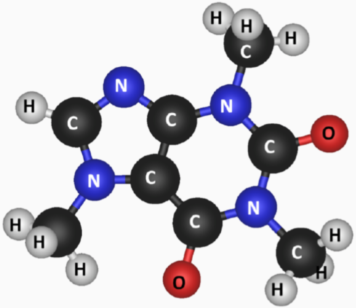
3D shape classification: results

SHREC'17 object retrieval competition metrics (perturbed micro-all)

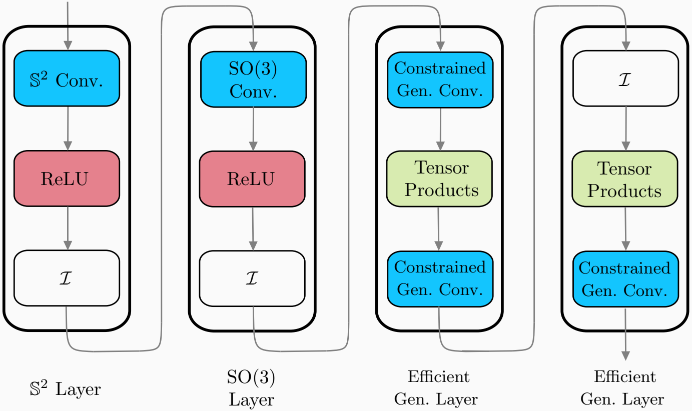
	P@N	R@N	F1@N	mAP	NDCG	Params
Kondor et al. 2018	0.707	0.722	0.701	0.683	0.756	>1M
Cohen et al. 2018	0.701	0.711	0.699	0.676	0.756	1.4M
Esteves et al. 2018	0.717	**0.737**	-	**0.685**	-	500k
Ours	**0.719**	0.710	**0.708**	0.679	**0.758**	**250k**

Atomization energy prediction: problem

Predict atomization energy of molecule give the atom charges and positions.



Atomization energy prediction: architecture



Atomization energy prediction: results

Test root mean squared (RMS) error for QM7 regression problem

	RMS	Params
Montavon et al. 2012	5.96	-
Cohen et al. 2018	8.47	1.4M
Kondor et al. 2018	7.97	>1.1M
Ours (MST)	**3.16**	337k
Ours (RMST)	3.46	**335k**

Summary

- Importance of encoding **equivariance to symmetry transforms** in order to capture fundamental physical understanding of generative process.

Summary

- Importance of encoding **equivariance to symmetry transforms** in order to capture fundamental physical understanding of generative process.
- Need for geometric deep learning techniques **constructed natively on manifolds**, such as the sphere.

Summary

- Importance of encoding **equivariance to symmetry transforms** in order to capture fundamental physical understanding of generative process.
- Need for geometric deep learning techniques **constructed natively on manifolds**, such as the sphere.
- Reviewed **spherical CNNs constructions**, with a focus on rotational equivariance (Cohen et al. 2018, Esteves et al. 2018, Kondor et al. 2018).

Summary

- Importance of encoding **equivariance to symmetry transforms** in order to capture fundamental physical understanding of generative process.
- Need for geometric deep learning techniques **constructed natively on manifolds**, such as the sphere.
- Reviewed **spherical CNNs constructions**, with a focus on rotational equivariance (Cohen et al. 2018, Esteves et al. 2018, Kondor et al. 2018).
- **Efficient generalised spherical CNNs** (Cobb et al. 2020; arXiv:2010.11661)
  - General framework that encompasses others as special cases.
  - Supports hybrid models to leverage strength of alternatives alongside each other.
  - New efficient layers to be used as primary building blocks.
  - State-of-the-art performance, both in terms of accuracy and parameter efficiency.

kagenova

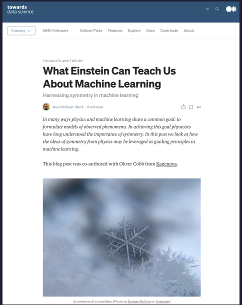
www.kagenova.com

Unlocking the potential of (geometric) deep learning to solve a wide range of problems in virtual reality (VR) and beyond.

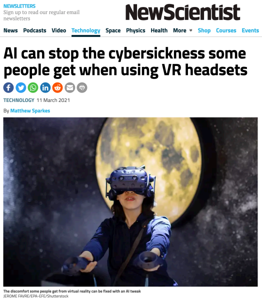
We're hiring!

hello@kagenova.com

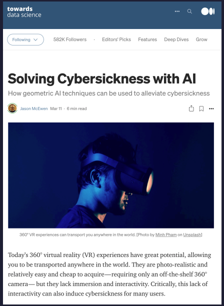
Read more in New Scientist, Digital Trends & Towards Data Science



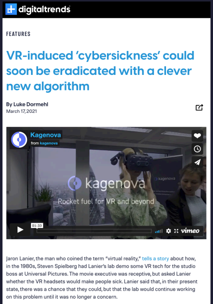
tinyurl.com/47nc6ccy



tinyurl.com/27zyc96b



tinyurl.com/2y7ybeyj



tinyurl.com/37pxkpen

Questions?