Equivariance I: Group Convolution

Gabriele Cesa

Senior Engineer, Qualcomm AI Research*

University of Amsterdam, AMLab

13/06/2024

*Qualcomm AI Research is an initiative of Qualcomm Technologies, Inc.

Ideal Outline

1. Group-Convolution
2. Steerable Convolution
3. Literature, Applications and more

Agenda

Equivariance

From Convolution to Group-Convolution

Some GCNNs in the literature

Theoretical Properties of GCNNs

Symmetries in data

satellite data

source: ESA, CC BY-SA 3.0

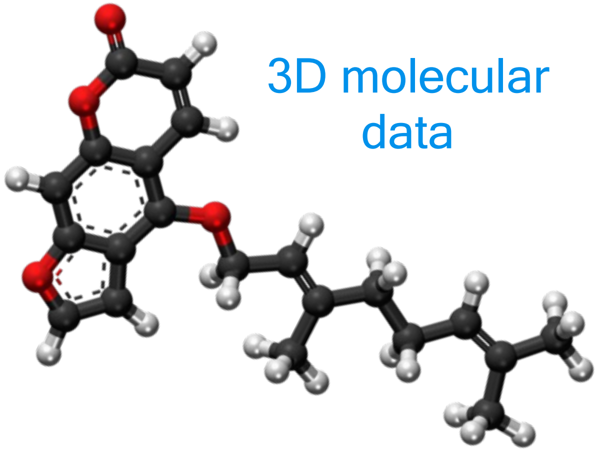
natural images

source: MikeLynch, CC BY-SA 3.0

translations

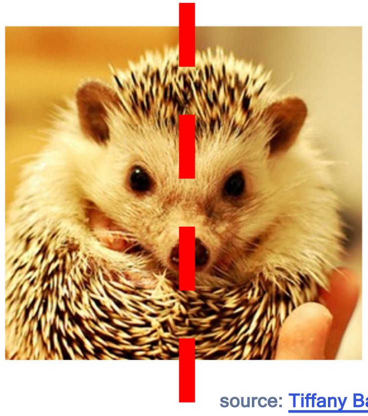
rotations

3D molecular data



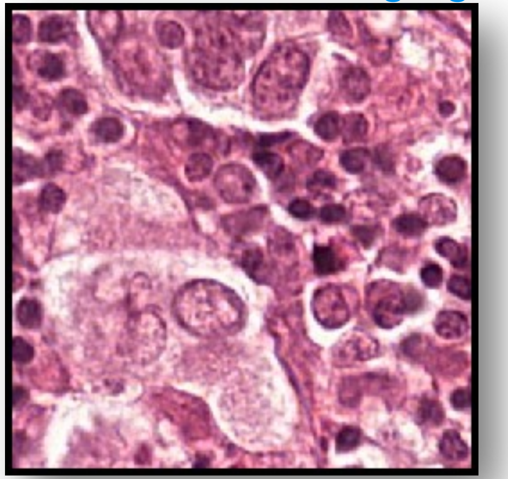
source: Jynto, CC0, Wikimedia Commons

medical imaging



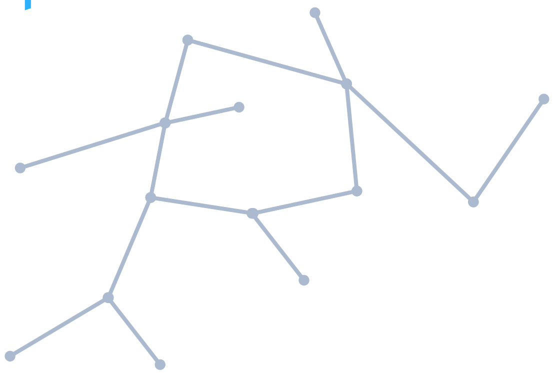
source: Tiffany Bailey, CC BY 2.0

reflections



source: CAMELYON17

permutations



Equivariance

input image

Neural network layer

$g.\Psi($  $) = \Psi(g.$  $)$

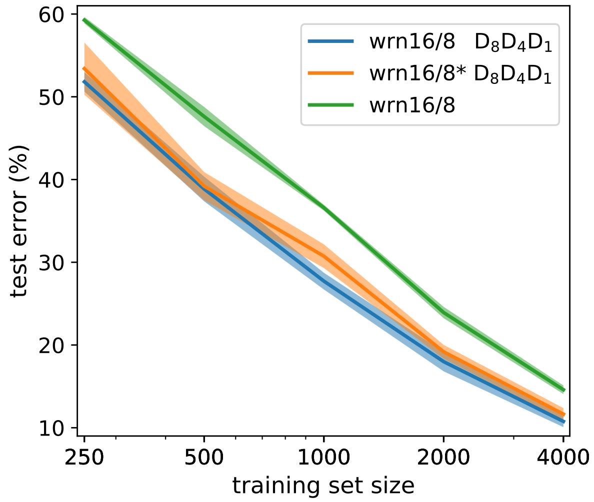
output transformation

input transformation

Equivariance: why?

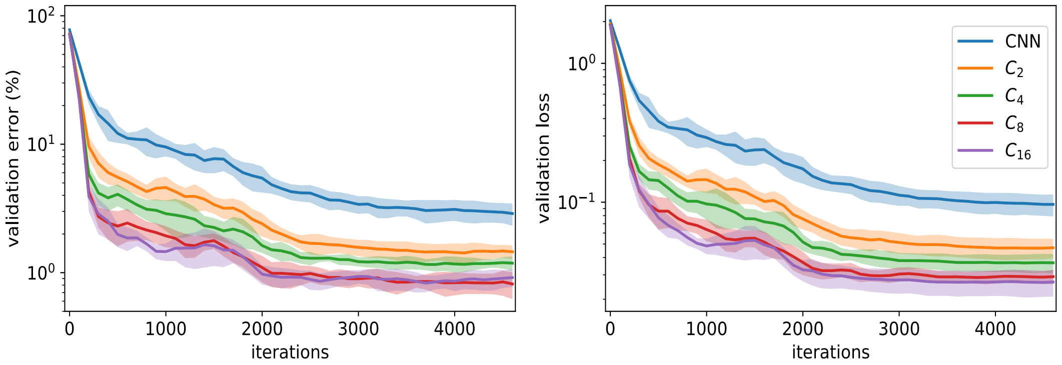
Data efficiency and improved generalization

- by shrinking the hypothesis space
- Automatically generalize over transformed versions of same input without observing them first



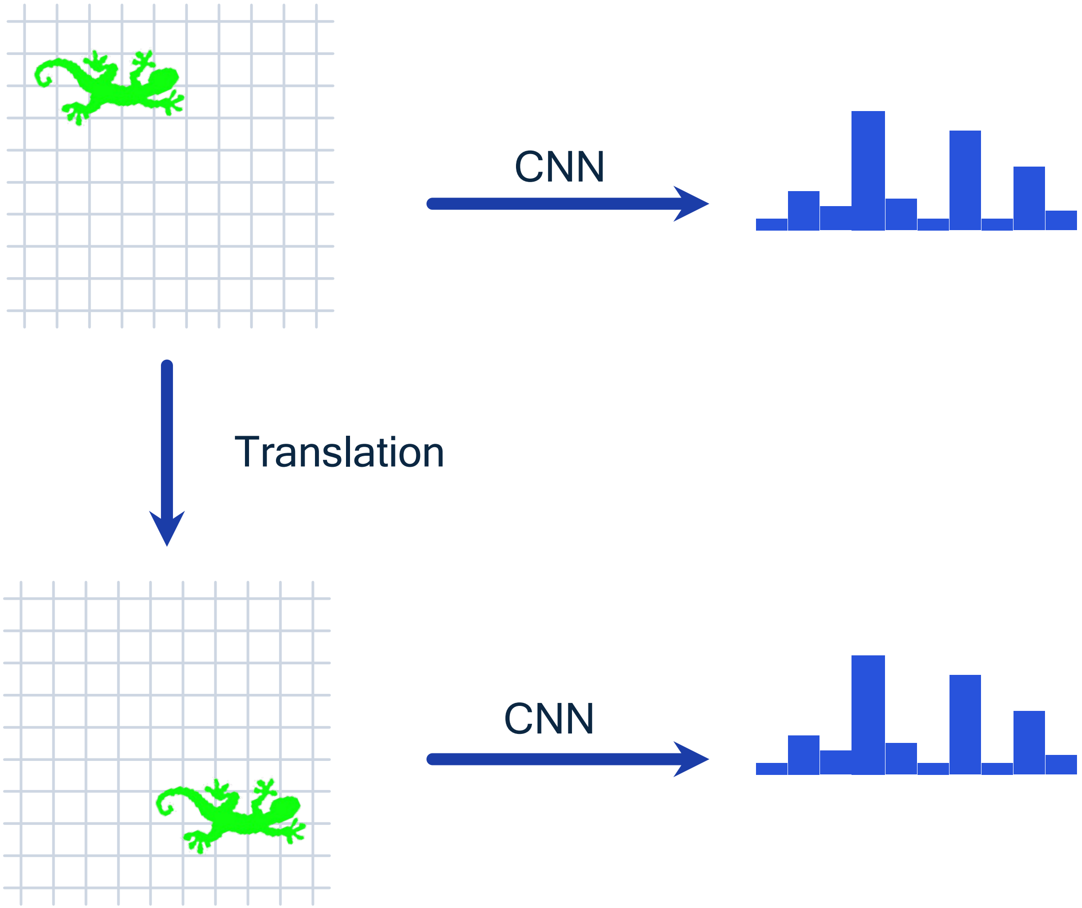
Generalization of equivariant and non-equivariant models on STL-10 dataset

Typically, faster convergence

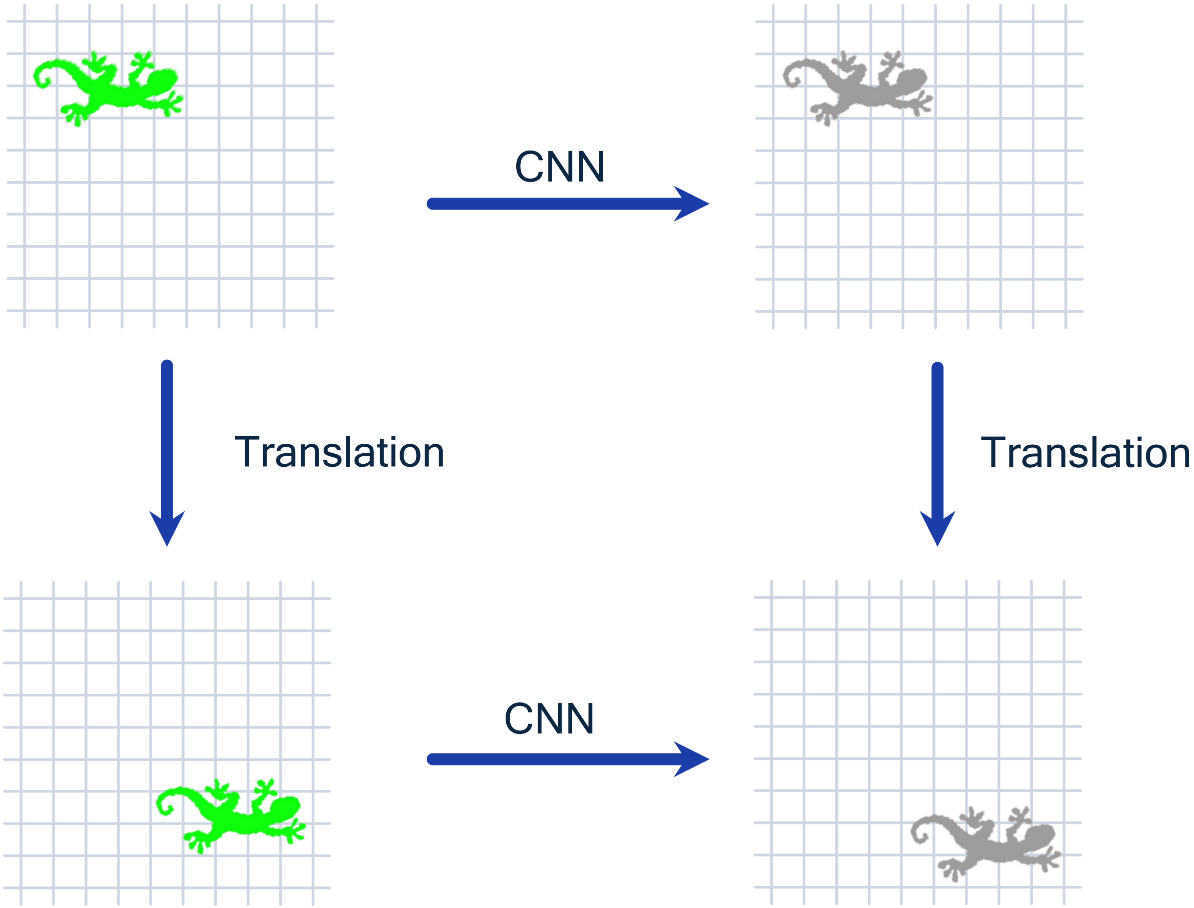


Convergence rate of different equivariant models evaluated on rotated MNIST dataset

Equivariance: CNN (translation invariance)



Equivariance: CNN (translation equivariance)

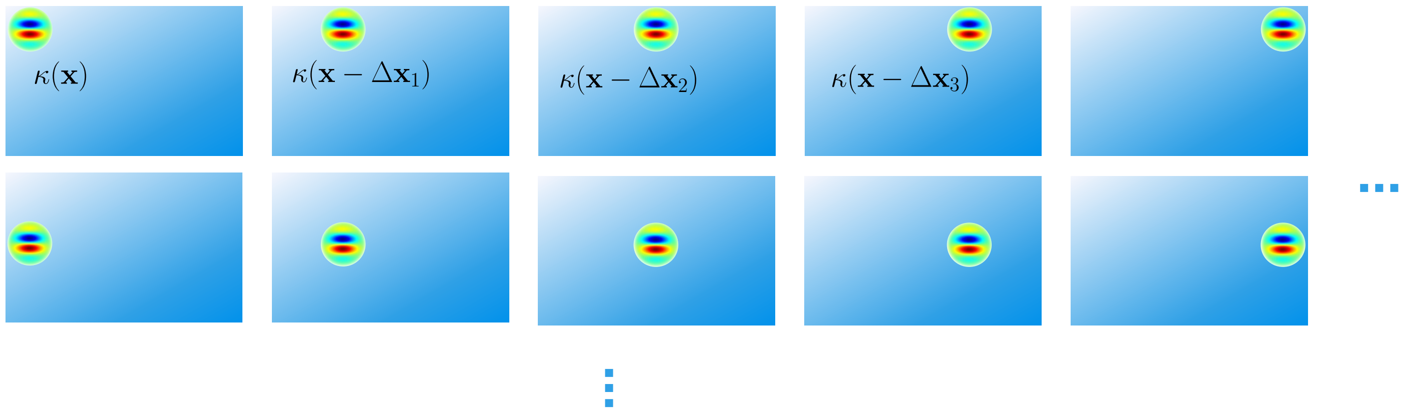


Convolution Layer

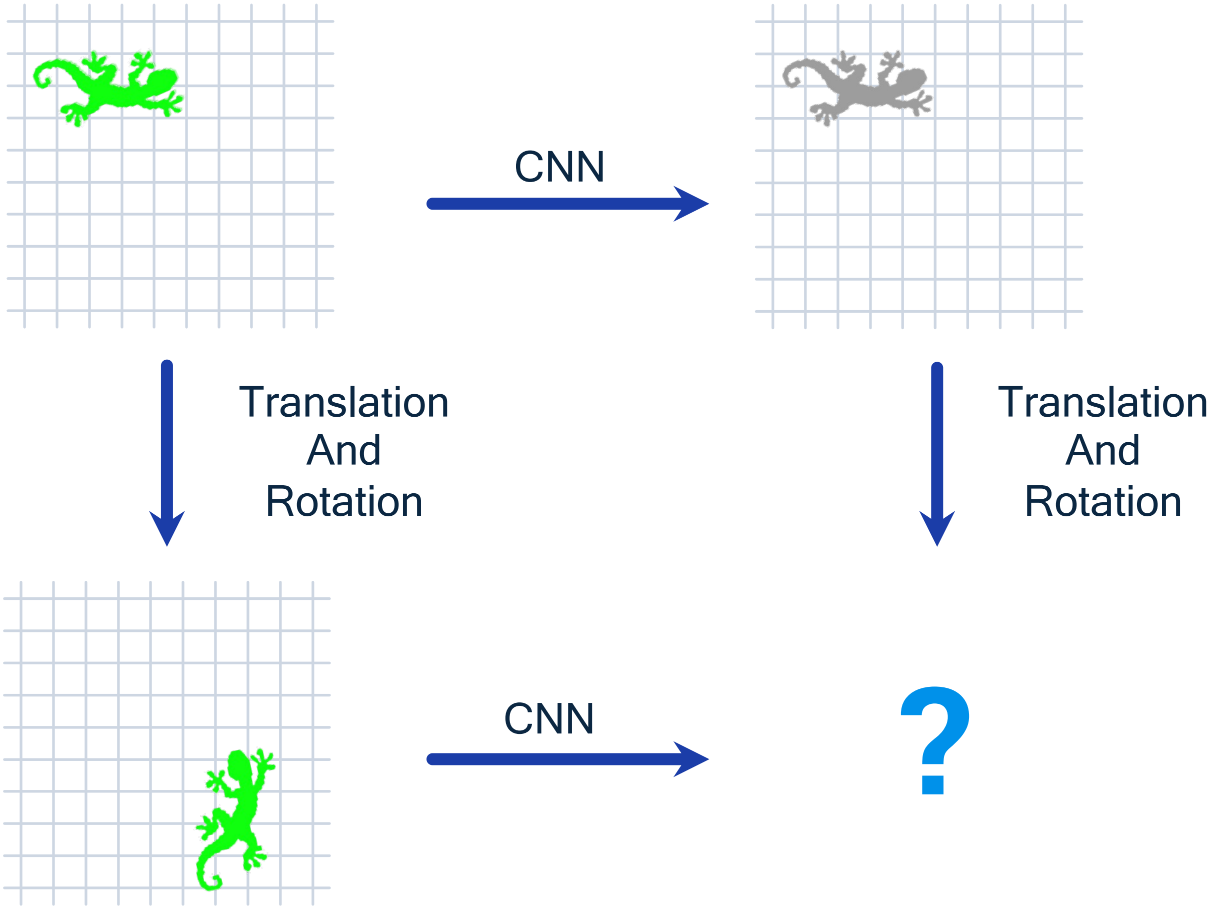
Actually a **cross-correlation**

$[\kappa \star f]: \mathbb{Z}^d \to \mathbb{R}^{c_{out}}$ $\kappa: \mathbb{Z}^d \to \mathbb{R}^{c_{out} \times c_{in}}$ $f: \mathbb{Z}^d \to \mathbb{R}^{c_{in}}$

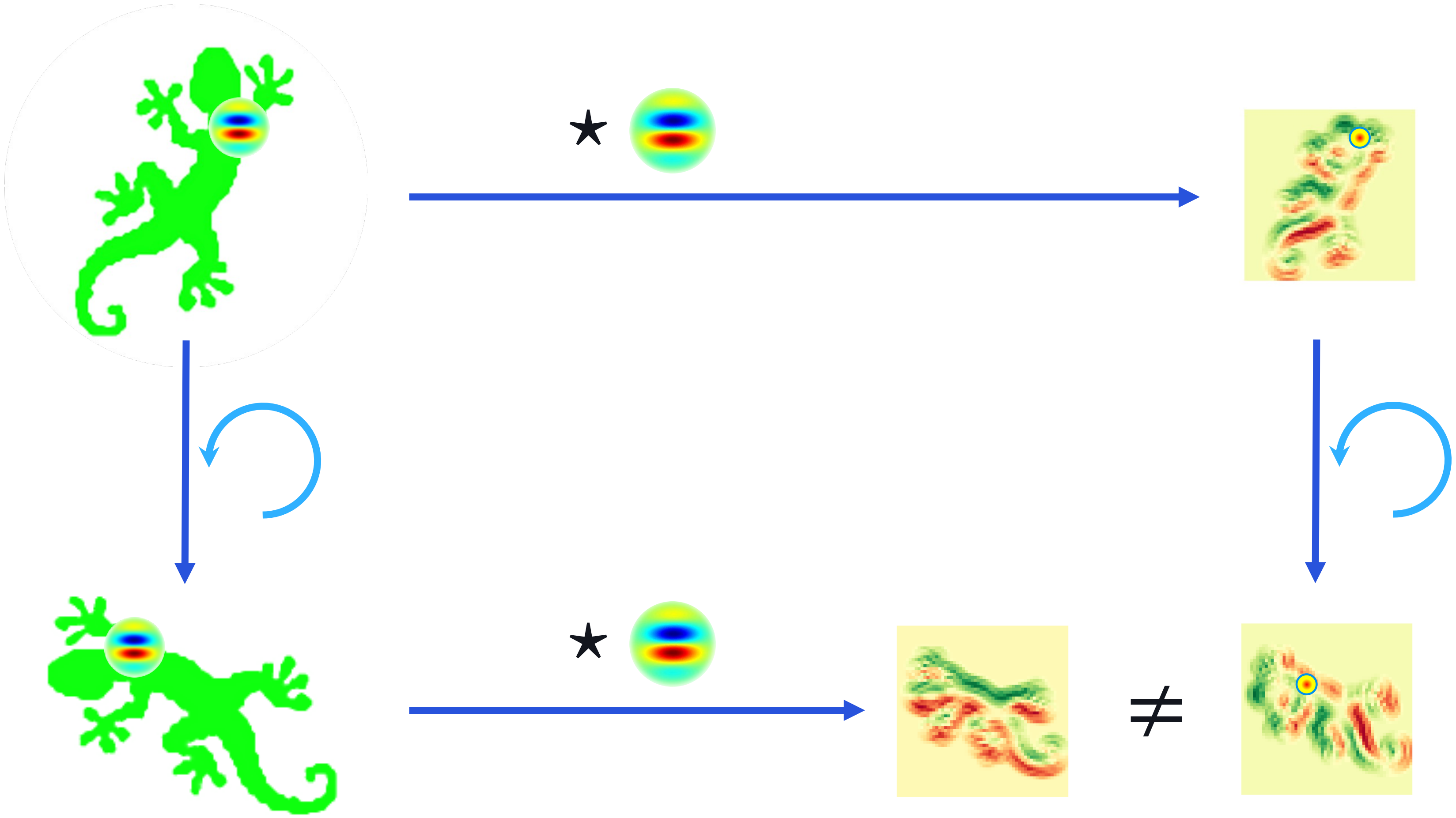
$$[\kappa \star f](\Delta \mathbf{x}) = \sum_{\mathbf{x} \in \mathbb{Z}^d} \kappa(\mathbf{x} - \Delta \mathbf{x}) \, f(\mathbf{x})$$



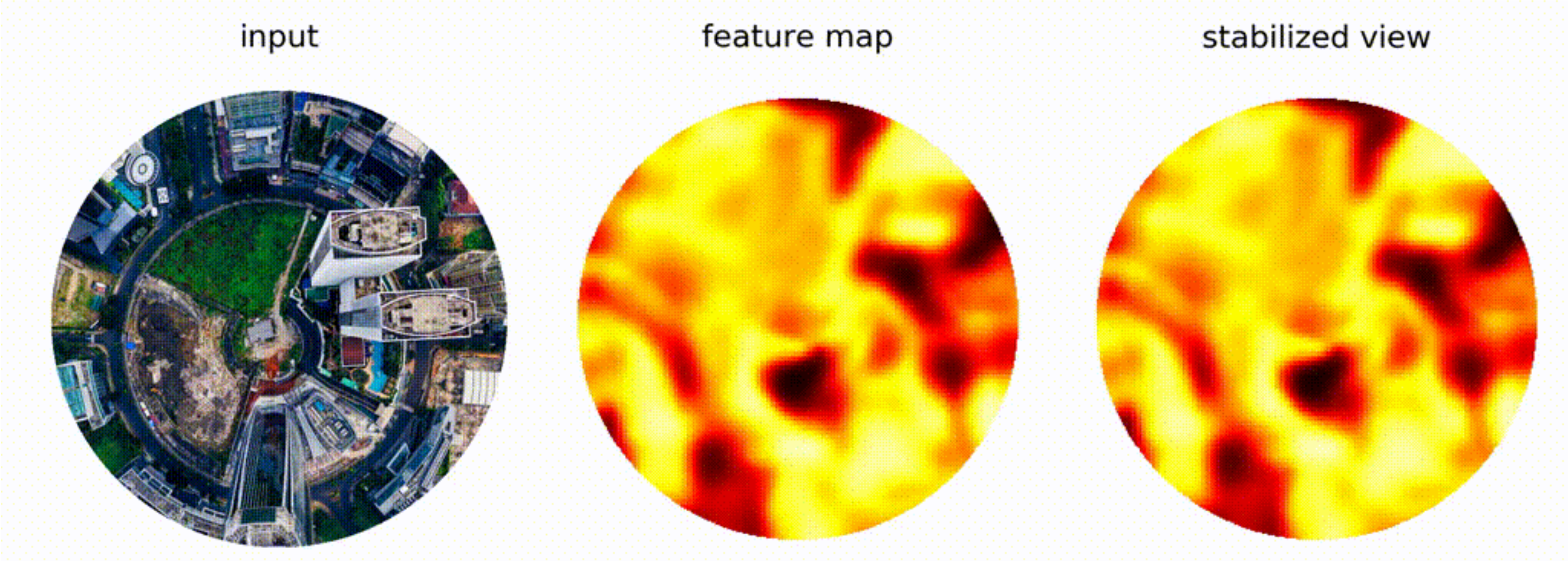
Equivariance: CNN (rotation equivariance?)



Equivariance: CNN (rotation equivariance?)

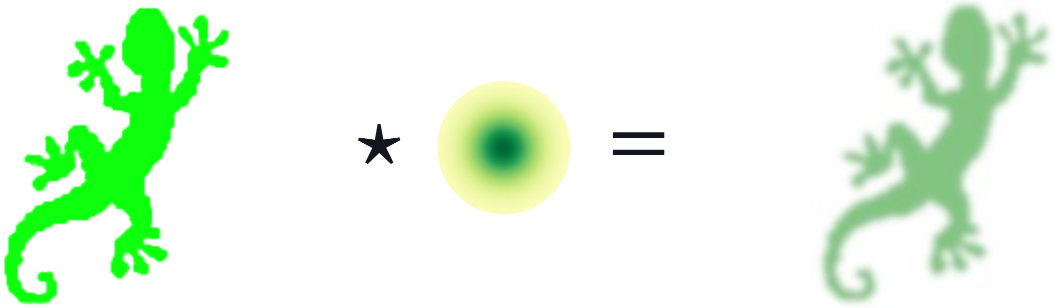


Equivariance: CNN (rotation equivariance?)

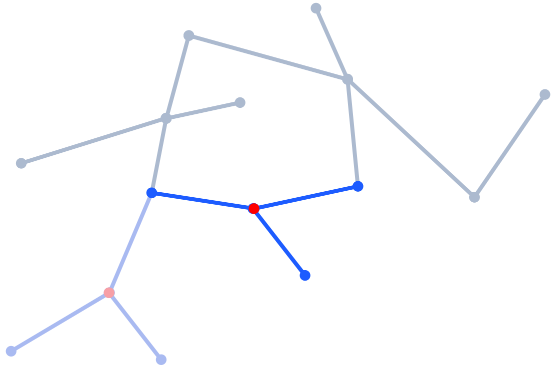


Isotropic Filters

- Isotropic filters: the response does not change when rotated



- Not very expressive
- Analogous to Graph Message Passing: no directional dependence!



Generalize Convolution (*template matching*)

Familiar **cross-correlation**

$[\kappa \star f]: \mathbb{Z}^d \to \mathbb{R}^{c_{out}}$ $\quad \kappa: \mathbb{Z}^d \to \mathbb{R}^{c_{out} \times c_{in}}$ $\quad f: \mathbb{Z}^d \to \mathbb{R}^{c_{in}}$

$$[\kappa \star f](\Delta \mathbf{x}) = \sum_{\mathbf{x} \in \mathbb{Z}^d} \kappa(\mathbf{x} - \Delta \mathbf{x})\ f(\mathbf{x})$$

Group cross-correlation

$[\kappa \star f]: G \to \mathbb{R}^{c_{out}}$ $\quad \kappa: X \to \mathbb{R}^{c_{out} \times c_{in}}$ $\quad f: X \to \mathbb{R}^{c_{in}}$

$$[\kappa \star f](g) = \sum_{\mathbf{x} \in X} \kappa(g^{-1}\mathbf{x})\ f(\mathbf{x})$$

g element of a **Group** G of transformations
(rotations, reflections, translations, …)

Taco S. Cohen and Max Welling, Group Equivariant Convolutional Networks,
International Conference on Machine Learning (ICML), 2016

Definition: **Group**

A group $(G,\cdot)$ is a *set* G and a *binary operation* $\cdot : G \times G \to G$ satisfying the following axioms:

- **associativity**: $\forall\, a, b, c \in G,\ a \cdot (b \cdot c) = (a \cdot b) \cdot c$
- **identity**: $\exists e \in G$ s.t. $\forall a \in G,\ a \cdot e = e \cdot a$
- **inverse**: $\forall g \in G,\ \exists!\ g^{-1} \in G$ s.t. $g g^{-1} = g^{-1} g = e$

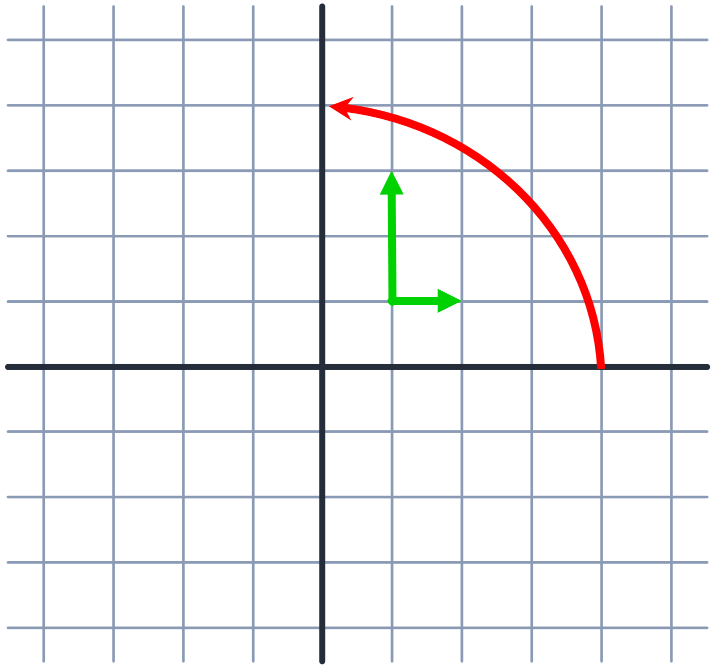
Relevant examples:
- discrete translations $(\mathbb{Z}^n, +)$
- continuous translations $(\mathbb{R}^n, +)$
- rotations $SO(n)$
- rotations and mirroring $O(n)$
- special isometries (rotations and translations) $SE(n) = (\mathbb{R}^n, +) \rtimes SO(n)$
- isometries (rotations, reflections, translations) $E(n) = (\mathbb{R}^n, +) \rtimes O(n)$
- permutation of N elements S_N

Running examples

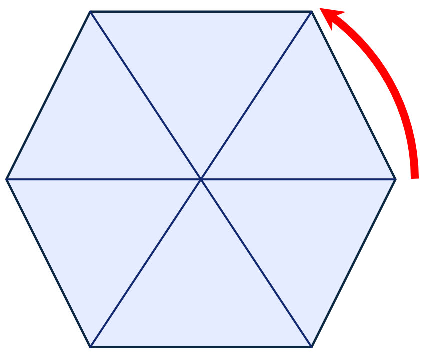
Discrete planar translations: $(\mathbb{Z}^2, +)$

Discrete planar rotations: C_n

Symmetries of squared grid: $\mathrm{p4} = (\mathbb{Z}^2, +) \rtimes C_4$



$\mathrm{p4} = (\mathbb{Z}^2, +) \rtimes C_4$



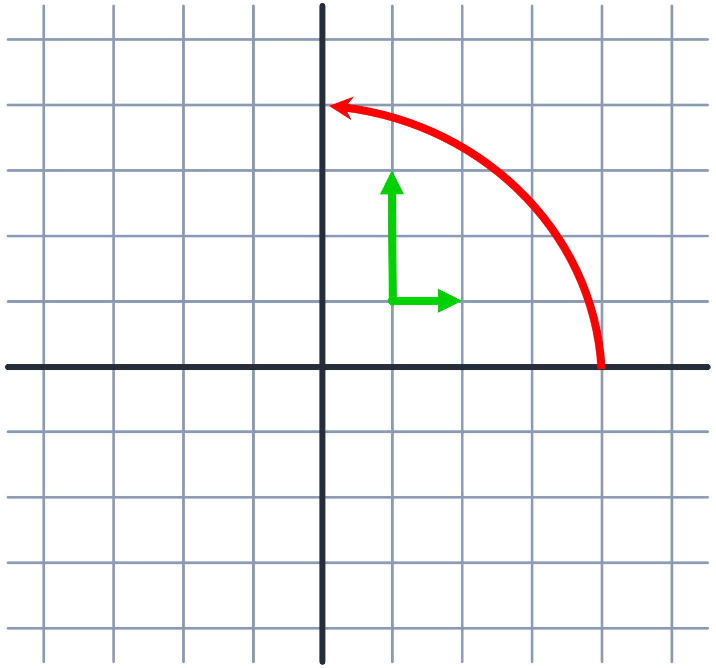
C_6

Running examples

Discrete planar translations: $(\mathbb{Z}^2, +)$

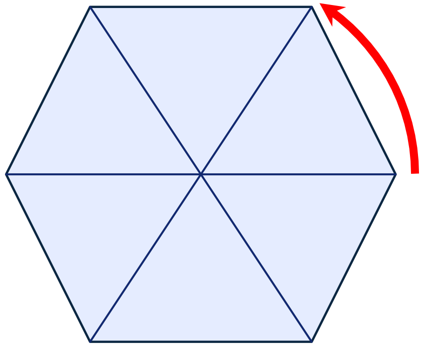
Discrete planar rotations: C_n

Symmetries of squared grid: $\mathrm{p4} = (\mathbb{Z}^2, +) \rtimes C_4$



$$\mathrm{p4} = (\mathbb{Z}^2, +) \rtimes C_4$$

If you don't know what this means,
for now, think of it as a
Cartesian Product $\times$
between two sets



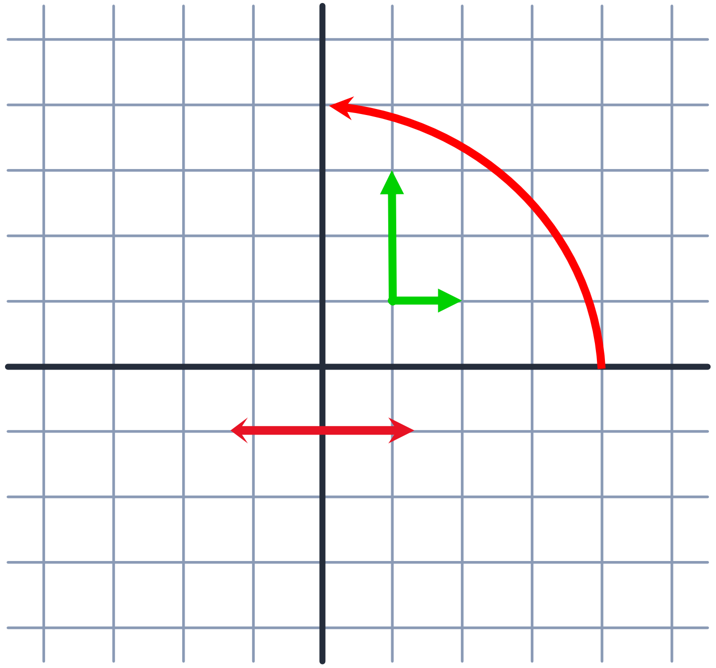
C_6

Running examples

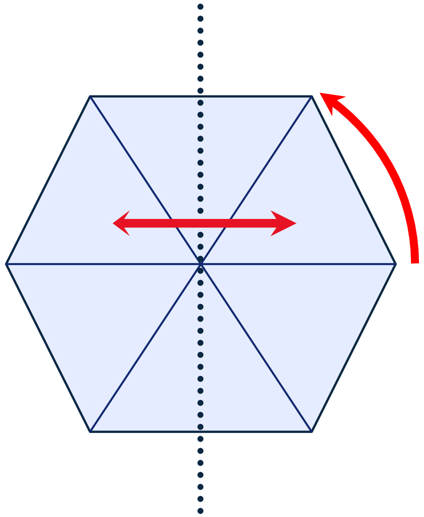
Discrete planar translations: $(\mathbb{Z}^2, +)$

Discrete planar rotations *and mirroring* : D_n

Symmetries of squared grid: $\mathrm{p4m} = (\mathbb{Z}^2, +) \rtimes D_4$



$\mathrm{p4m} = (\mathbb{Z}^2, +) \rtimes D_4$



D_6

Definition: **Group Action**

Given a group G, a ***G*-space** X is a set equipped with a

group action $.: G \times X \to X$, i.e. a map satisfying the following axioms

- **identity**: $\forall x \in X, \quad x = e.x$
- **compatibility**: $\forall x \in X, \ \forall a, b \in G \quad a.(b.x) = (a \cdot b).x$

Relevant examples:

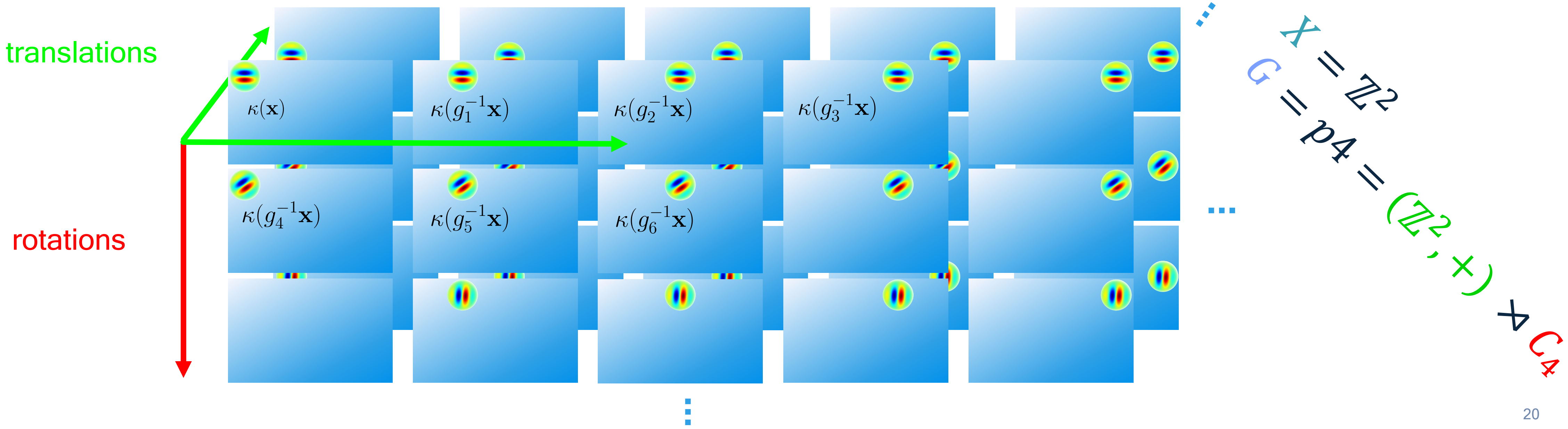
- discrete translations $G = (\mathbb{Z}^2, +)$ over a plane $X = \mathbb{R}^2$
- rotations $G = SO(3)$ of a sphere $X = \mathcal{S}^2$
- isometries $G = E(2)$ of a plane $X = \mathbb{R}^2$
- Discrete isometries $p4m$ of grid $X = \mathbb{Z}^2$
- permutation $G = S_N$ of a set of N ordered objects $X = (x_1, x_2, \dots, x_N)$
- continuous translations $G = (\mathbb{R}^2, +)$ over a plane $X = \mathbb{R}^2$

Generalize Convolution

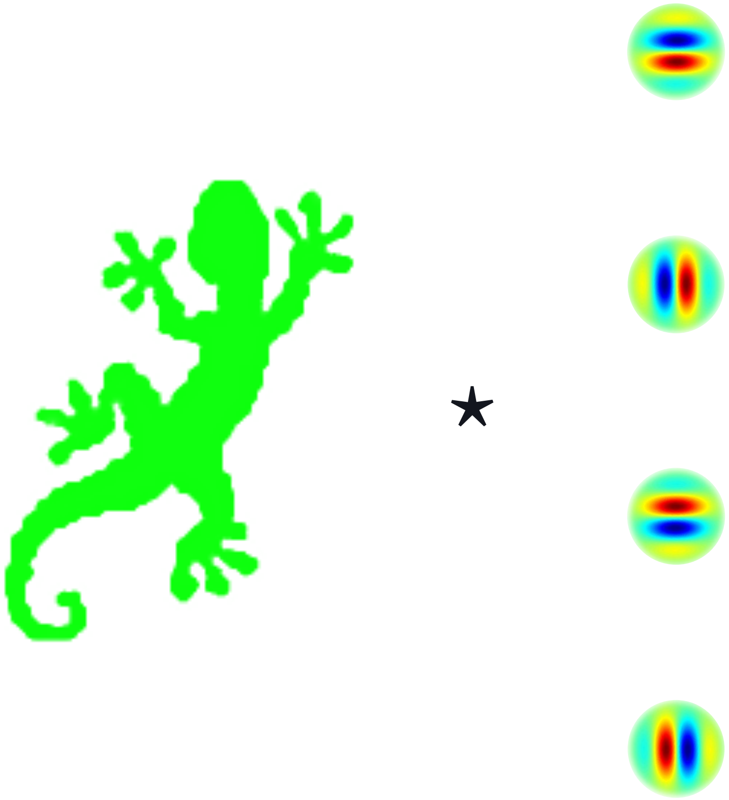
Group cross-correlation (**lifting convolution**):

$[\kappa \star f]: G \to \mathbb{R}^{c_{out}}$ $\quad \kappa: X \to \mathbb{R}^{c_{out} \times c_{in}}$ $\quad f: X \to \mathbb{R}^{c_{in}}$

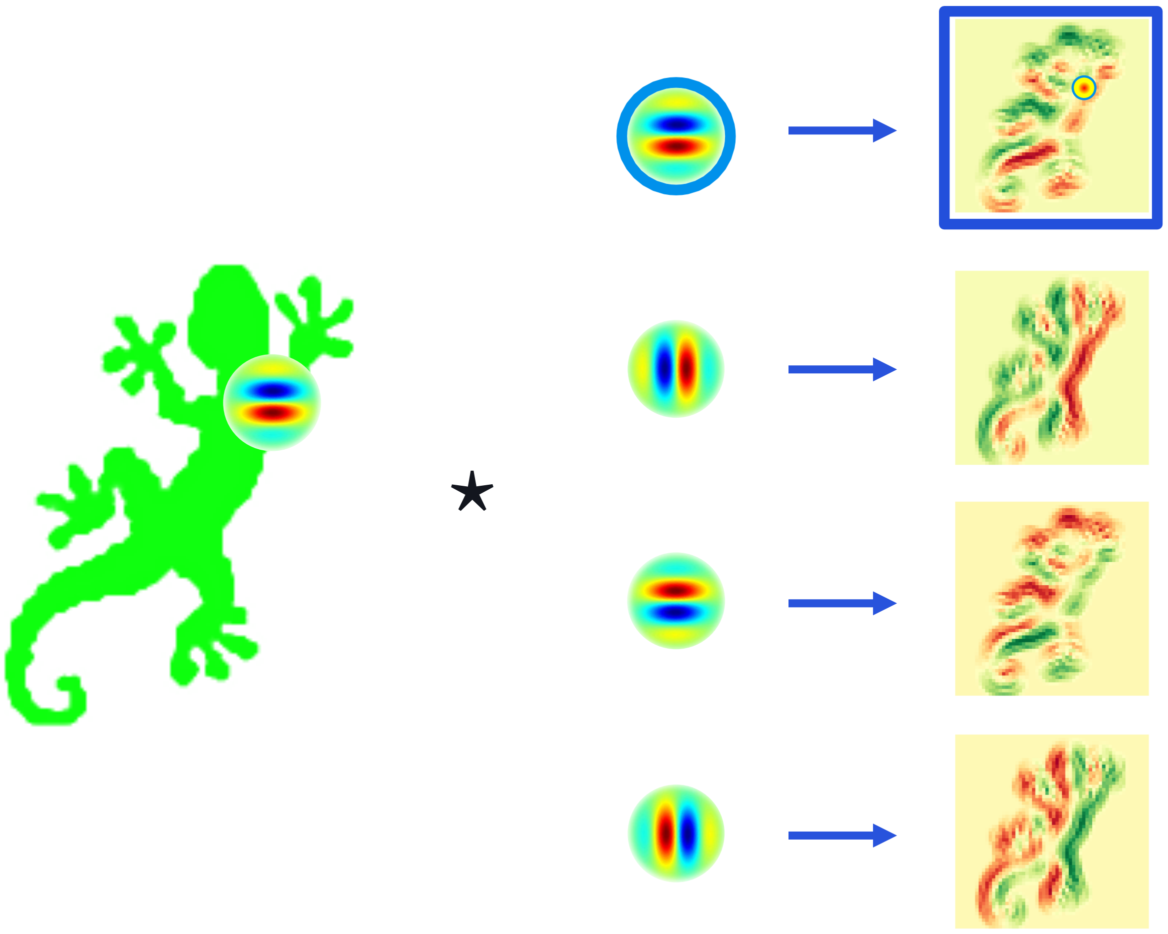
$$[\kappa \star f](g) = \sum_{\mathbf{x} \in X} \kappa(g^{-1}\mathbf{x})\ f(\mathbf{x})$$



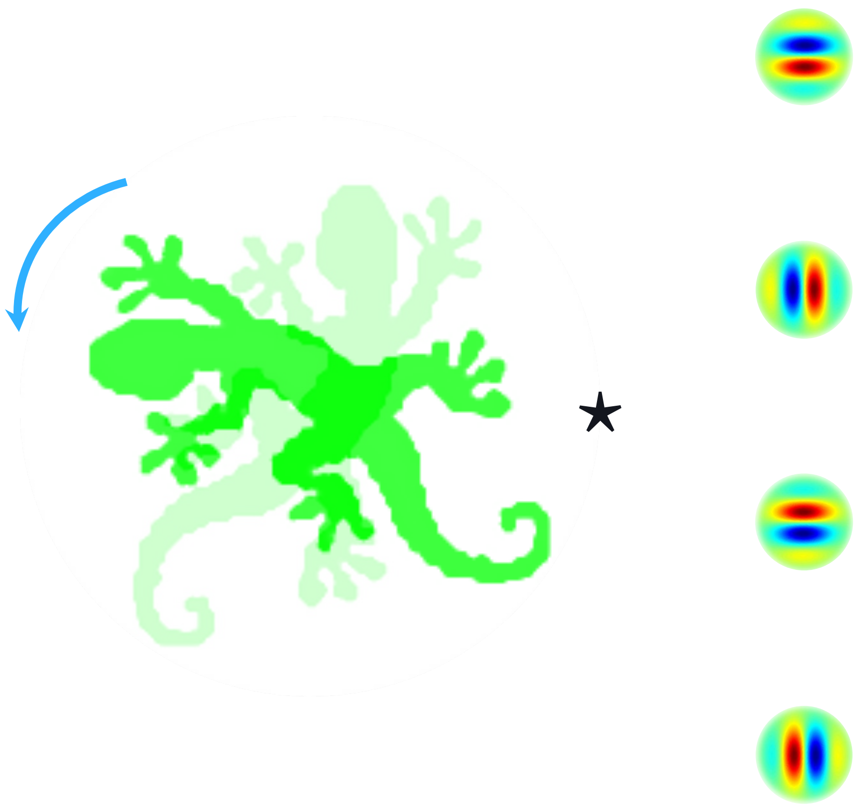
Group Convolution



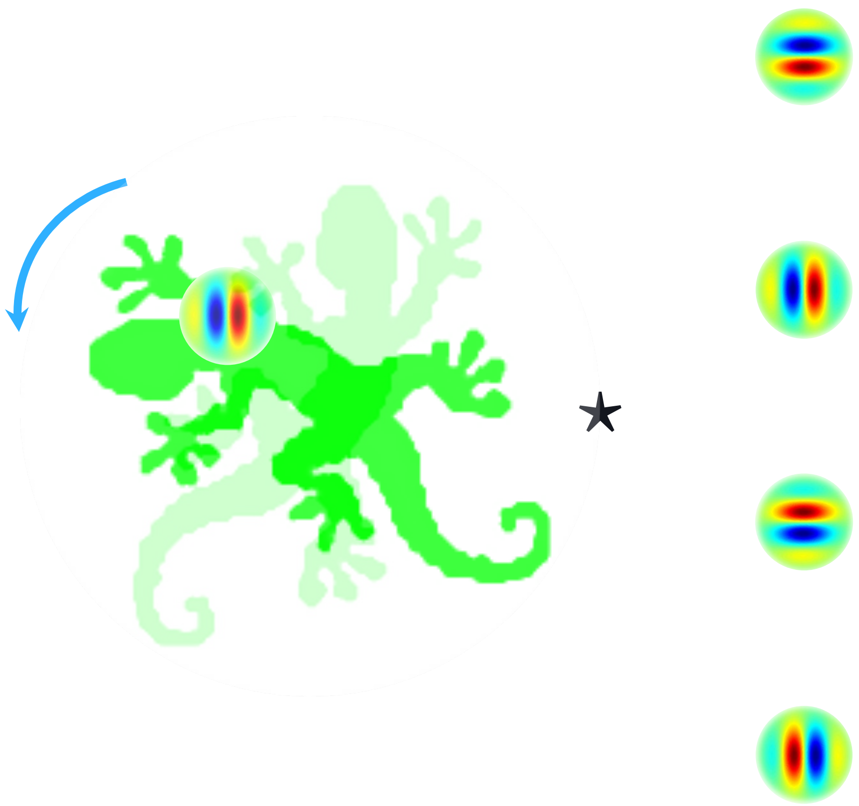
Group Convolution



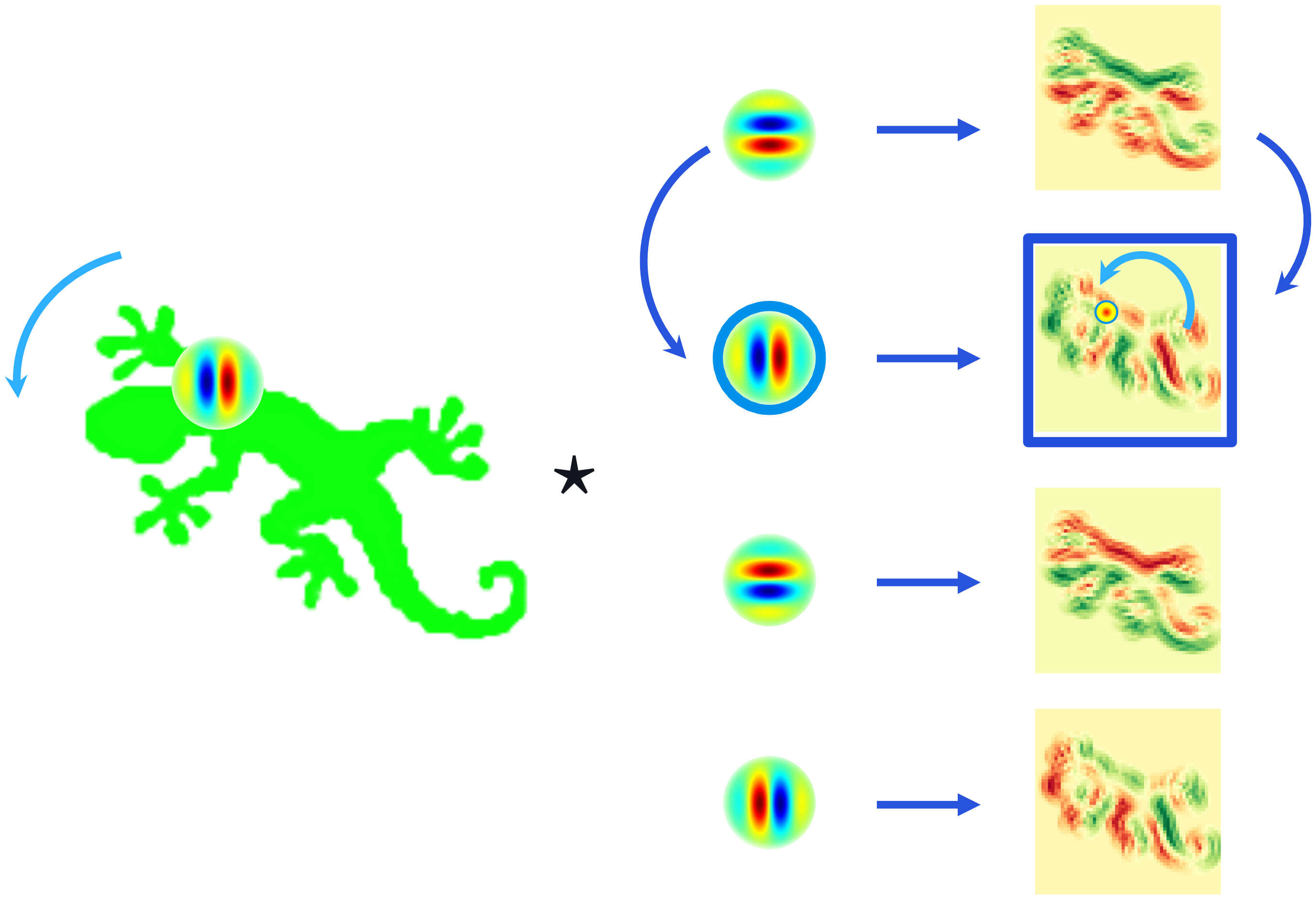
Group Convolution



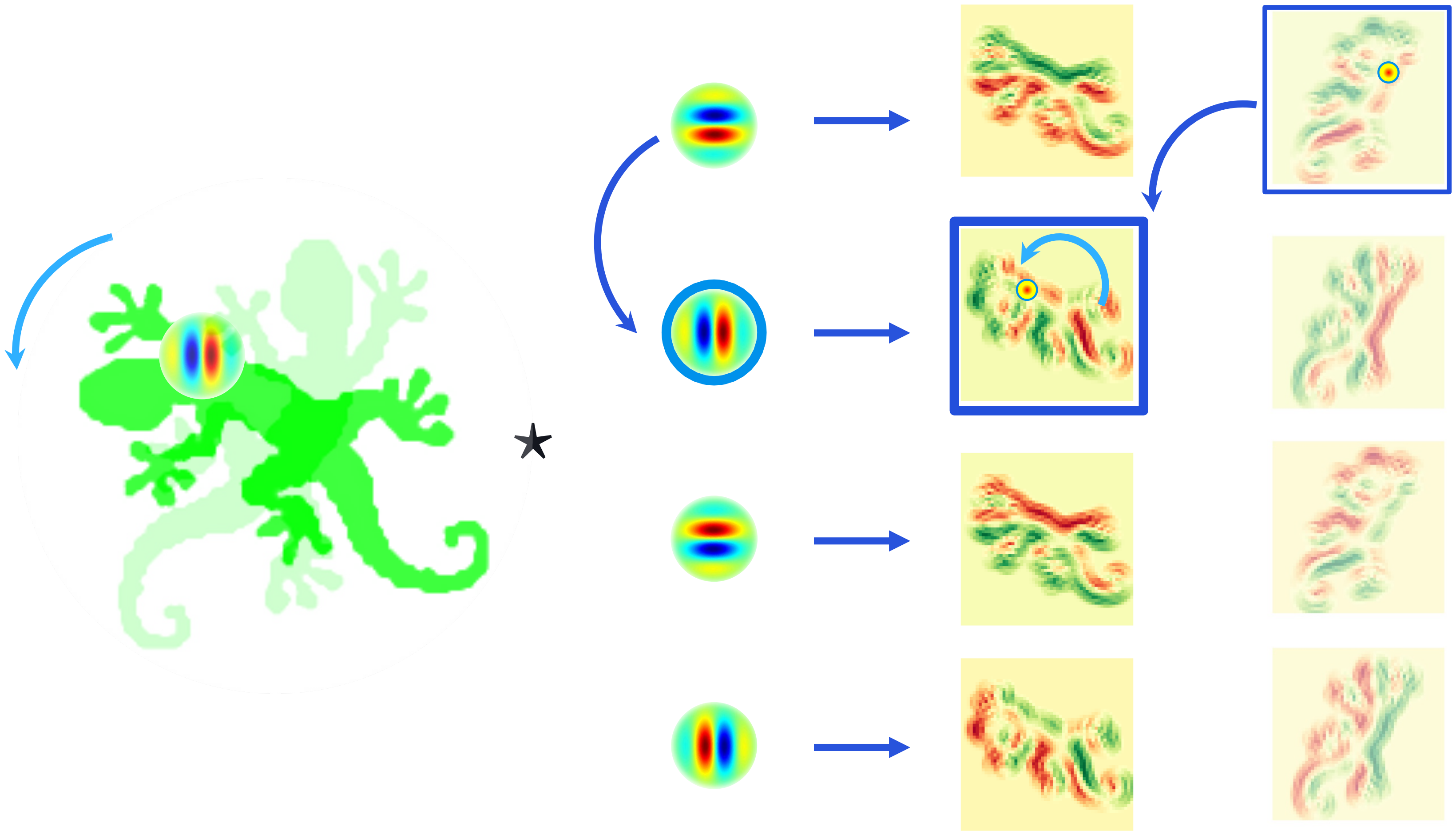
Group Convolution



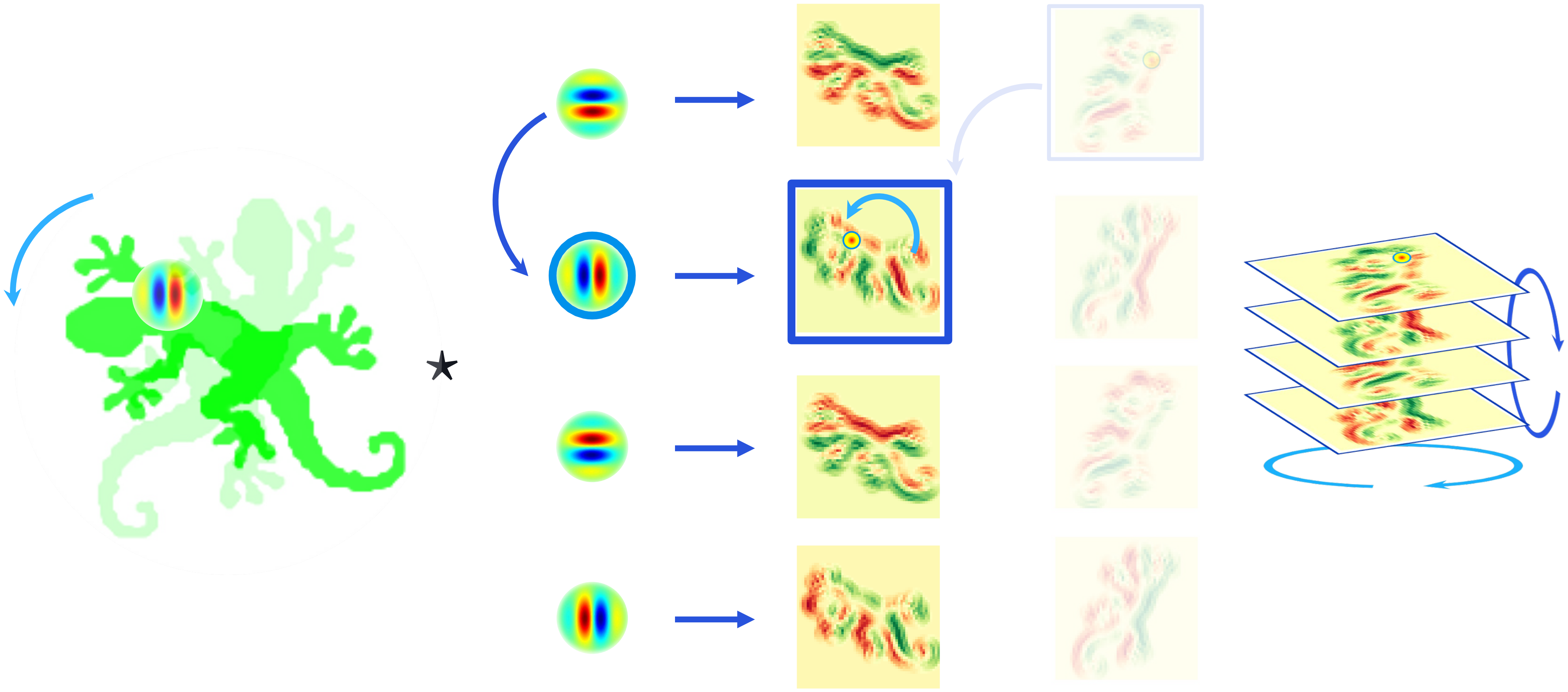
Group Convolution



Group Convolution



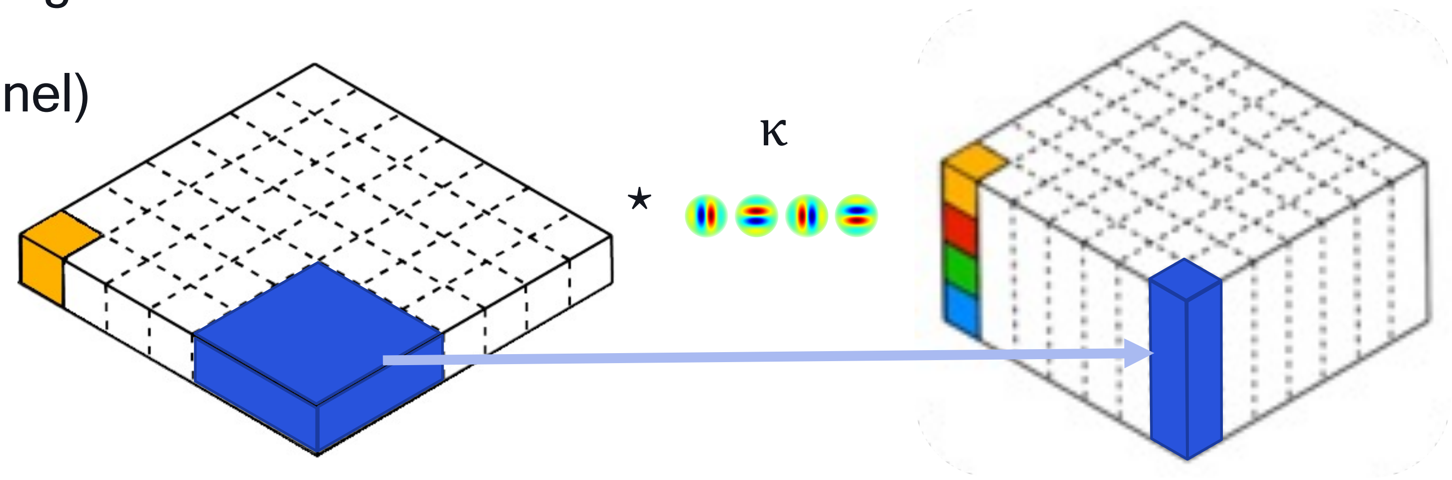
Group Convolution



Group Convolution (**lifting convolution**)

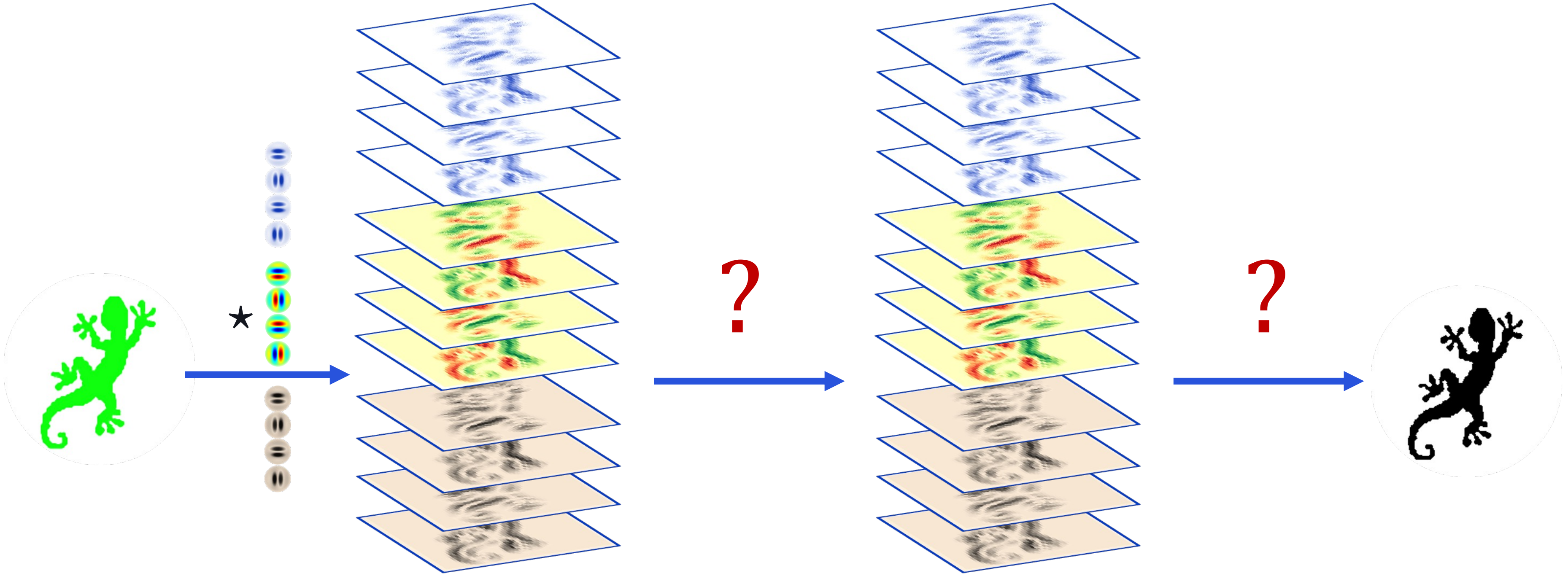
Input image

(1 channel)



Group convolution
applies rotated versions
of the same filter

Deep Neural Network



$f: \mathbb{Z}^n \rightarrow \mathbb{R}^{c_{in}}$ $f: G \rightarrow \mathbb{R}^c$ $f: G \rightarrow \mathbb{R}^c$ $f: \mathbb{Z}^n \rightarrow \mathbb{R}^{c_{out}}$

Definition: **Group Action**

Given a group G, a G-space X is a set equipped with a

group action $.: G \times X \to X$, i.e. a map satisfying the following axioms

- **identity**: $\forall x \in X, \quad x = e.x$
- **compatibility**: $\forall x \in X, \ \forall a, b \in G \quad a.(b.x) = (a \cdot b).x$

Relevant examples:

- discrete translations $G = (\mathbb{Z}^2, +)$ over a plane $X = \mathbb{R}^2$
- rotations $G = SO(3)$ of a sphere $X = \mathcal{S}^2$
- isometries $G = E(2)$ of a plane $X = \mathbb{R}^2$
- Discrete isometries $p4m$ of grid $X = \mathbb{Z}^2$
- permutation $G = S_N$ of a set of N ordered objects $X = (x_1, x_2, \dots, x_N)$
- continuous translations $G = (\mathbb{R}^2, +)$ over a plane $X = \mathbb{R}^2$
- a group G acting on its own elements $X = G$

Generalize Convolution

Group cross-correlation (intermidiate layers):

$[\kappa \star f]: \boldsymbol{G} \to \mathbb{R}^{c_{out}}$ $\kappa: \boldsymbol{G} \to \mathbb{R}^{c_{out} \times c_{in}}$ $f: \boldsymbol{G} \to \mathbb{R}^{c_{in}}$

$$[\kappa \star f](g) = \sum_{h \in \boldsymbol{G}} \kappa(g^{-1} h) \; f(h)$$

Let's get formal

Let G be a group and X a G-space

Denote the space of *square-integrable functions* on X as $L^2(X)$

$$f : X \to \mathbb{R}$$

$$f \in L^2(X) \iff \int_{x \in X} |f(\boldsymbol{x})|^2 \, d\mu(\boldsymbol{x}) < \infty$$

e.g. μ Haar measure when $X = G$

Let's get formal

Let G be a group and X a G-space

Denote the space of *square-integrable functions* on X as $L^2(X)$

$L^2(X)$ is a ***vector space*** with *inner product* :

$$\langle f_1, f_2 \rangle := \int_{x \in X} f_1(x) f_2(x) \, d\mu(\boldsymbol{x})$$

Let's get formal

Let G be a group and X a G-space

Denote the space of square-integrable functions on X as $L^2(X)$

$L^2(X)$ carries an **action** of the group G

$$g: L^2(X) \to L^2(X), \qquad f \mapsto g.f$$

$$[g.f](\boldsymbol{x}) \coloneqq f(g^{-1}\,\boldsymbol{x})$$

Let’s get formal

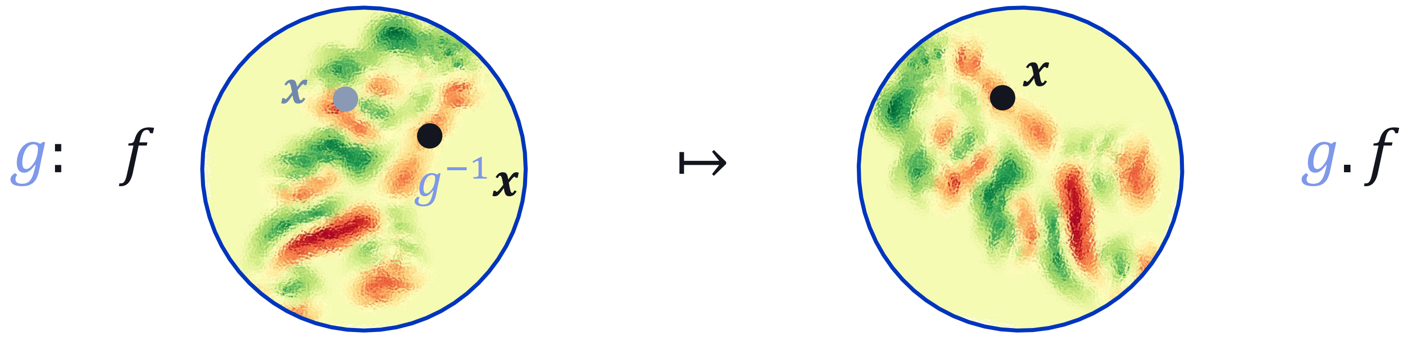
Let G be a group and X a G-space

Denote the space of square-integrable functions on X as $L^2(X)$

$L^2(X)$ carries an **action** of the group G

$$g: L^2(X) \to L^2(X), \qquad f \mapsto g.f$$

$$[g.f](\boldsymbol{x}) \coloneqq f(g^{-1}\,\boldsymbol{x})$$



Let's get formal

Let G be a group and X a G-space

Denote the space of square-integrable functions on X as $L^2(X)$

We can express the convolution as

$$[\kappa \star f](g) = \int_{x \in X} \kappa(g^{-1}\,\boldsymbol{x})\ f(\boldsymbol{x}) d\mu(x)$$

$$= \langle g.\kappa, f \rangle$$

$$[g.\kappa](\boldsymbol{x}) := \kappa(g^{-1}\,\boldsymbol{x})$$

Let's get formal

Let G be a group and X a G-space

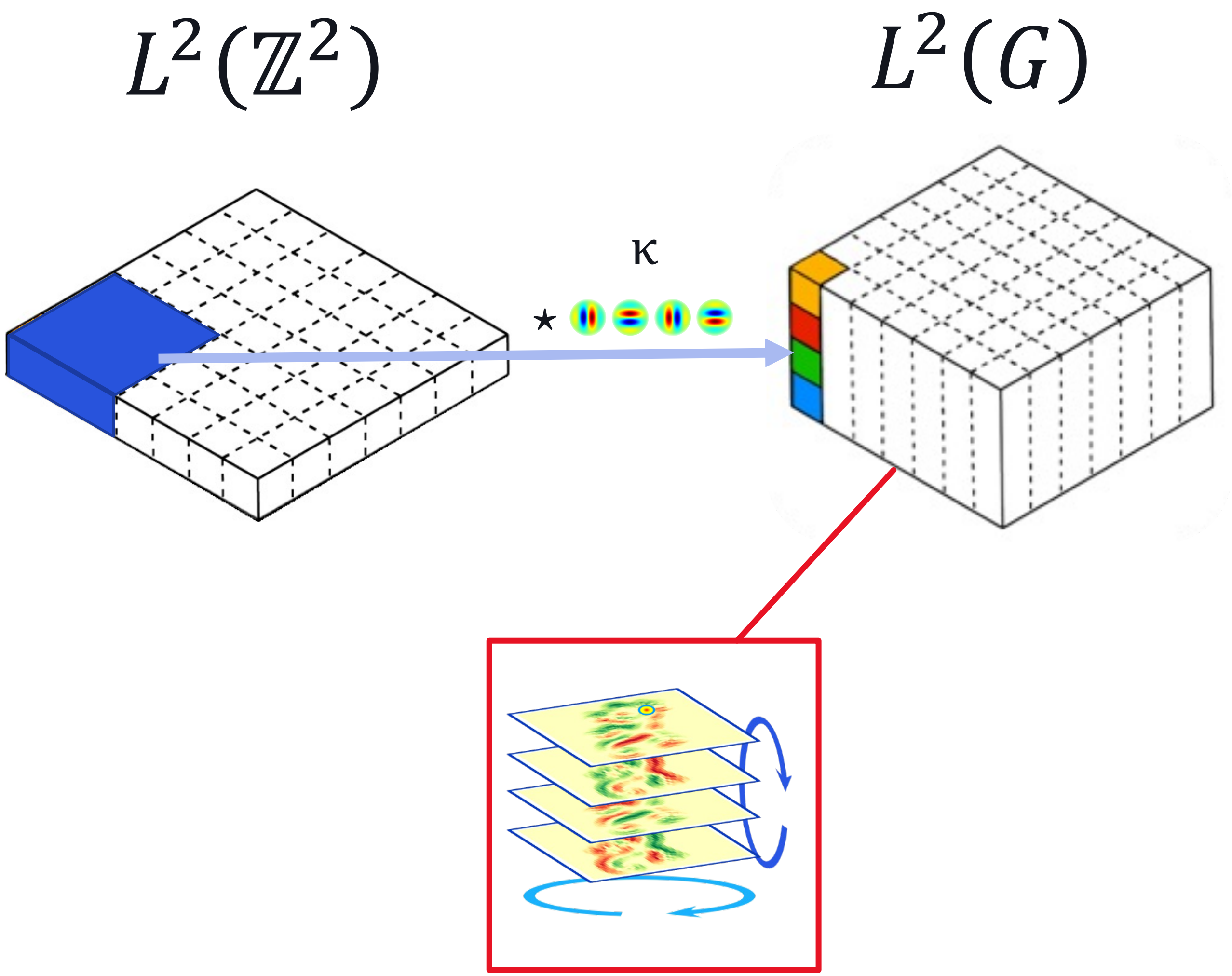
Denote the space of square-integrable functions on X as $L^2(X)$

We can express the convolution as

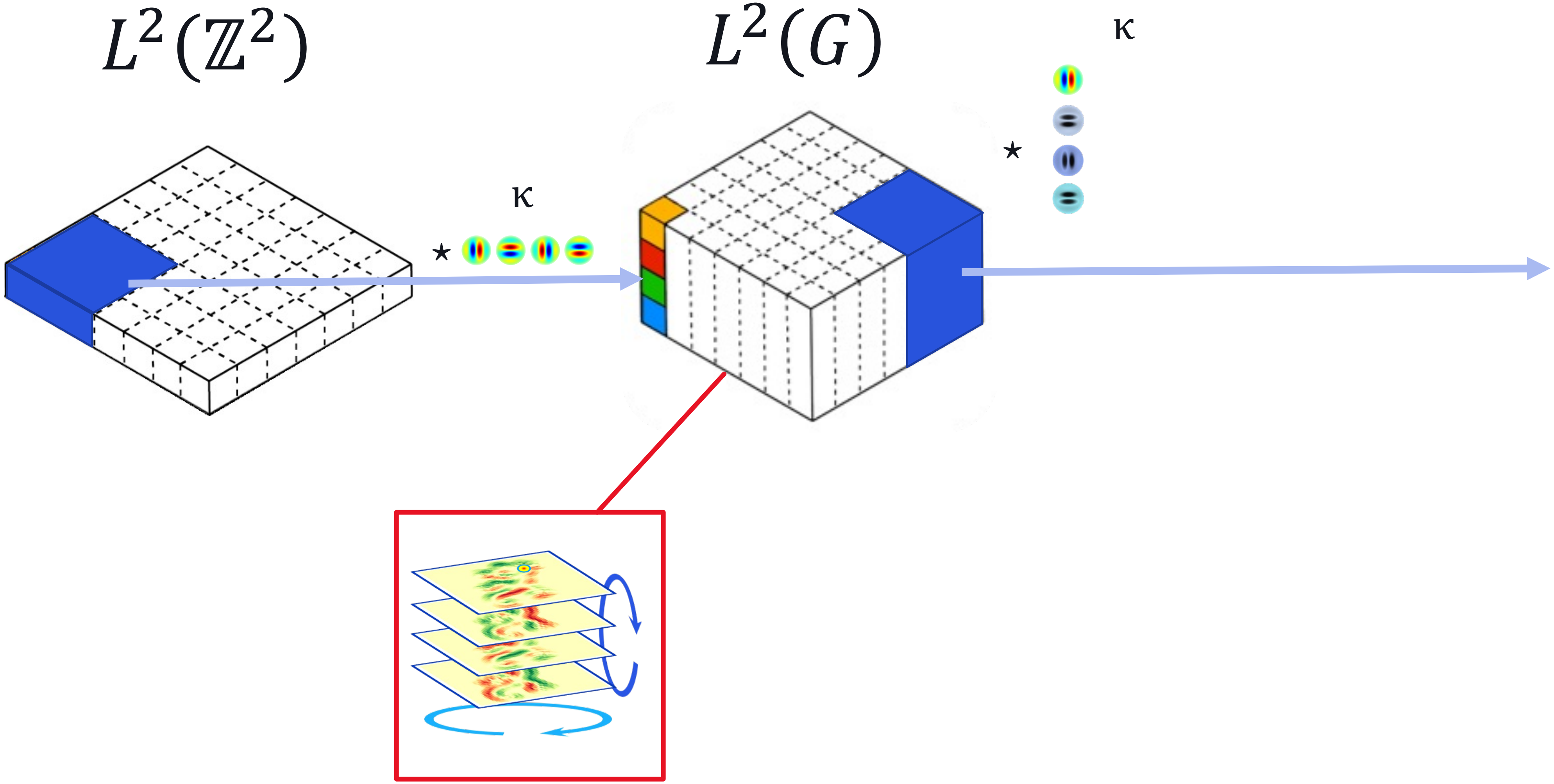
$$[\kappa \star f](g) = \int_{x \in X} \kappa(g^{-1}\,\boldsymbol{x})\ f(\boldsymbol{x}) d\mu(x) = \langle g.\kappa, f \rangle$$

$$[\kappa \star \cdot]: L^2(X) \to L^2(G)$$

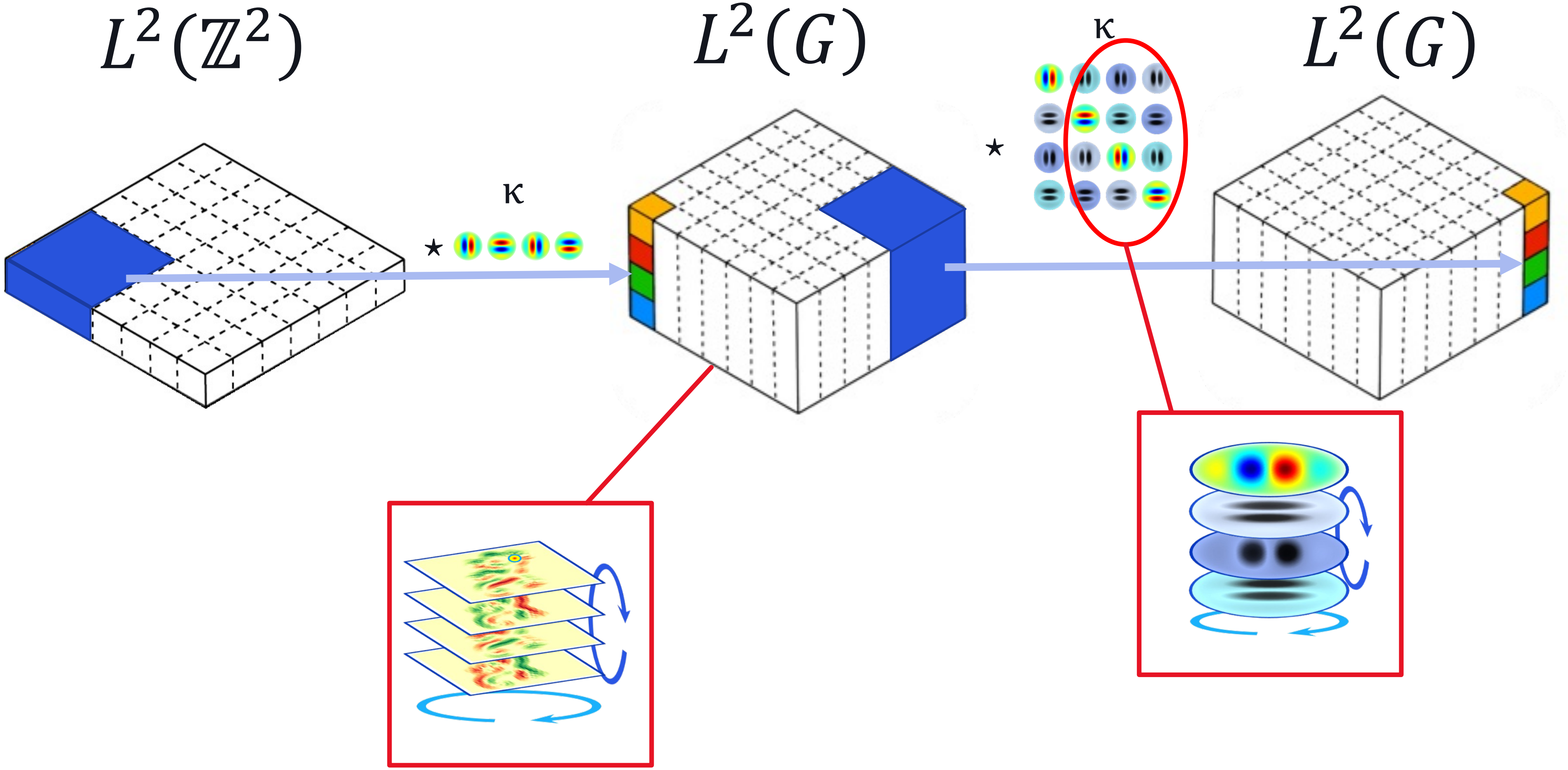
Group Convolution



Group Convolution



Group Convolution



Exercise

Show that group convolution is **equivariant**

$$g.[\kappa \star f] = [\kappa \star g.f]$$

Using:

$$[\kappa \star f](g) = \int_{x \in X} \kappa(g^{-1}\,\boldsymbol{x})\; f(\boldsymbol{x}) d\mu(x)$$

$$[g.f](\boldsymbol{x}) \coloneqq f(g^{-1}\,\boldsymbol{x})$$

Let's get formal - side note

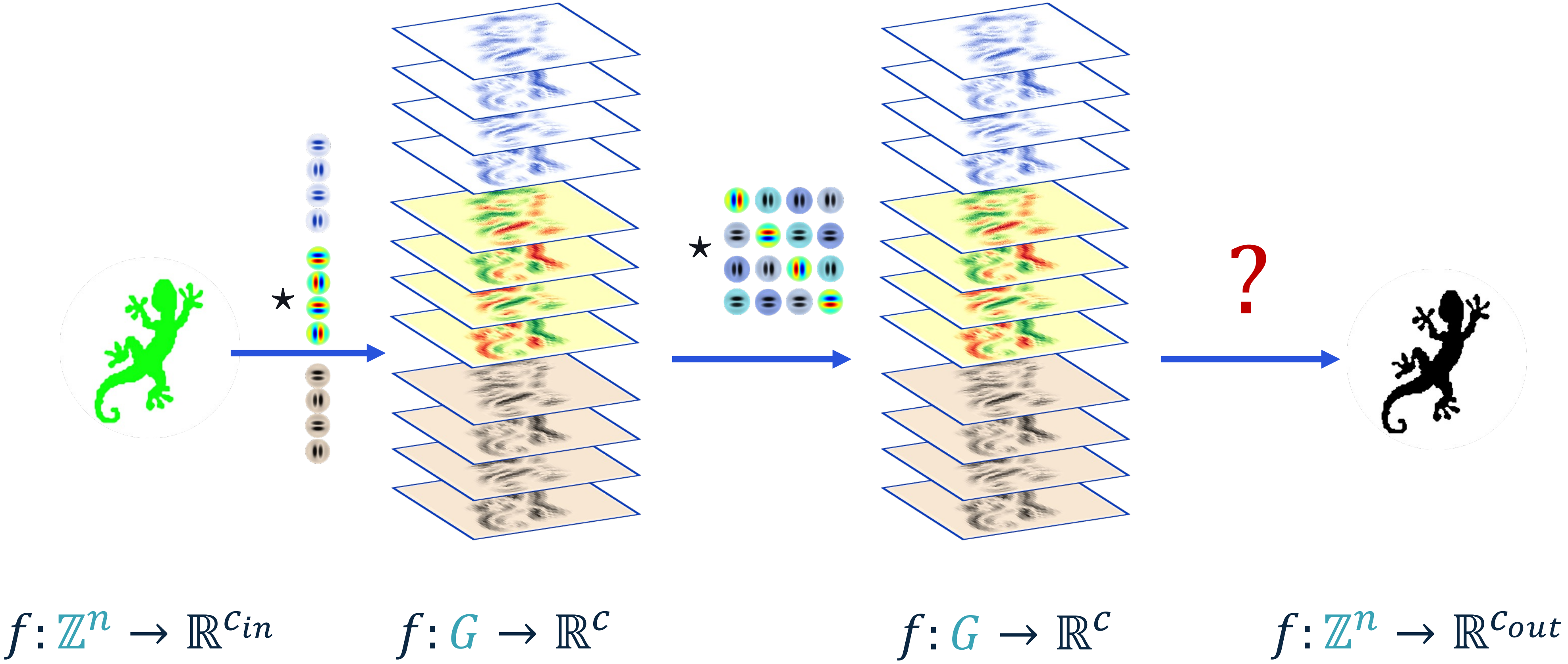
Let G be a group and V a vector space carrying an action of G

We can express the convolution as

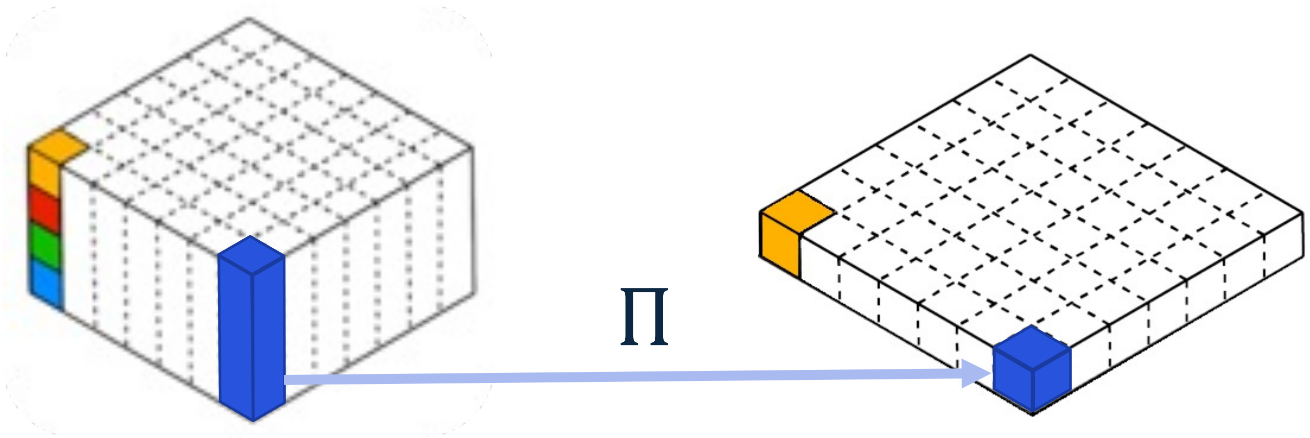
$$[\kappa \star f](g) = \langle g.\kappa, f \rangle$$

$$[\kappa \star \cdot]: V \to L^2(G)$$

Deep Neural Network



Pooling Layer

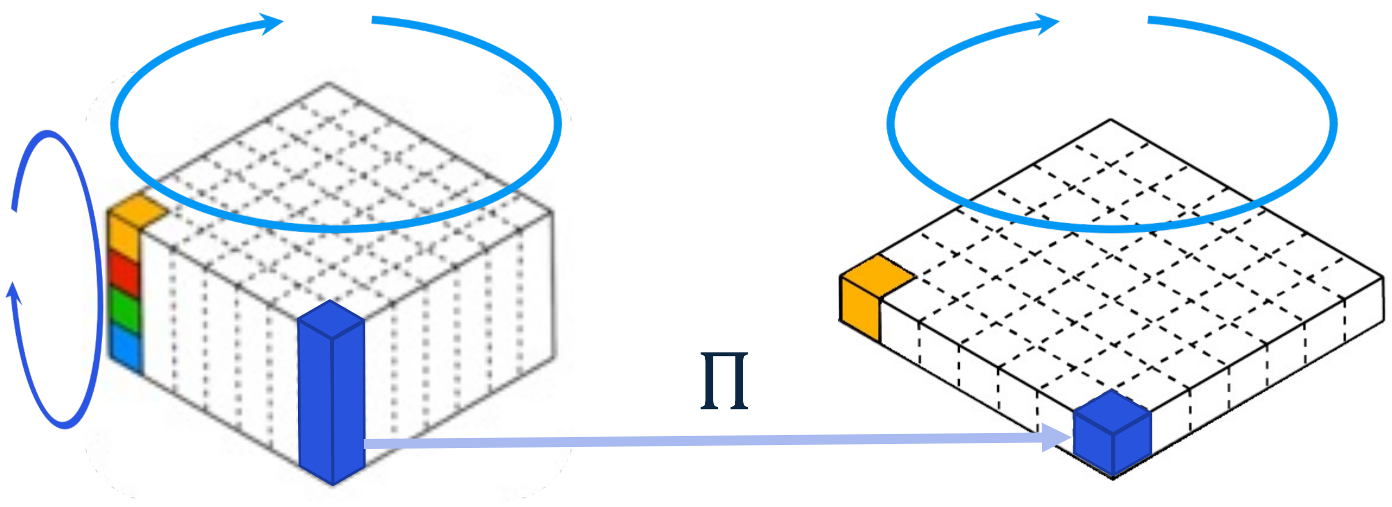


Pooling operation $\mathbf{\Pi}$, for example

- Max pooling
- Average pooling

Is the output *rotation invariant*?

Pooling Layer

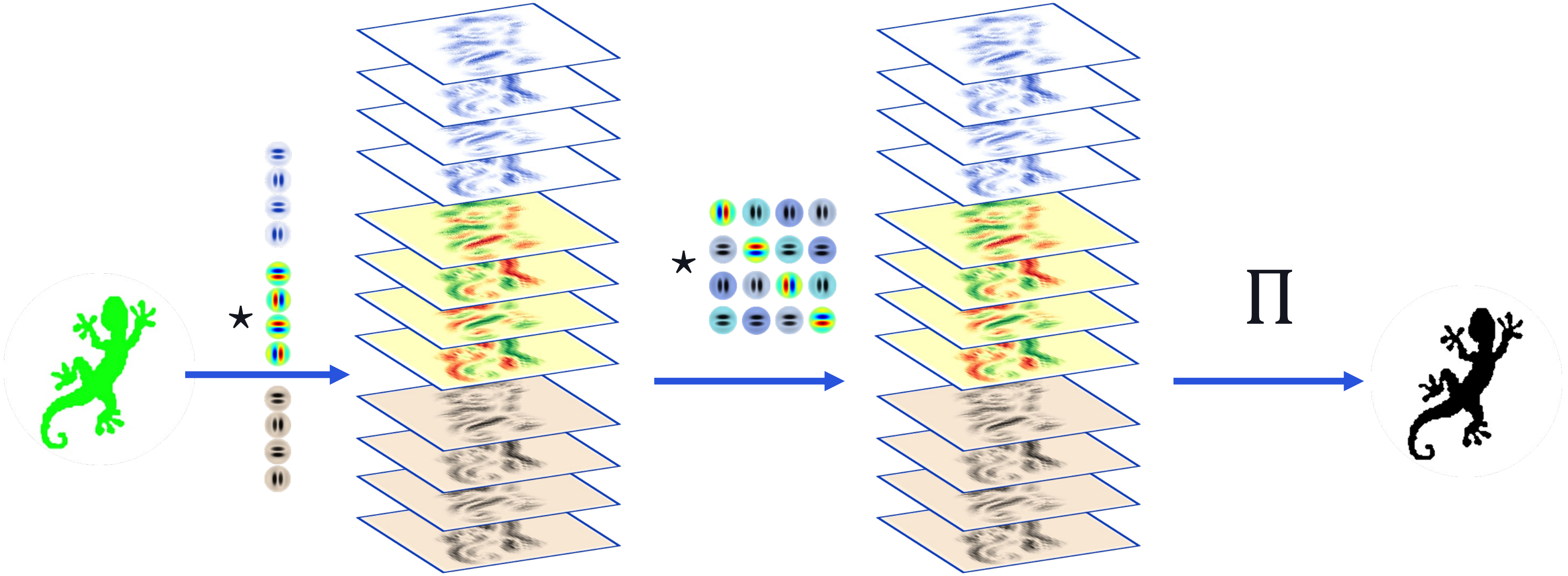


Pooling operation $\mathbf{\Pi}$, for example

- Max pooling
- Average pooling

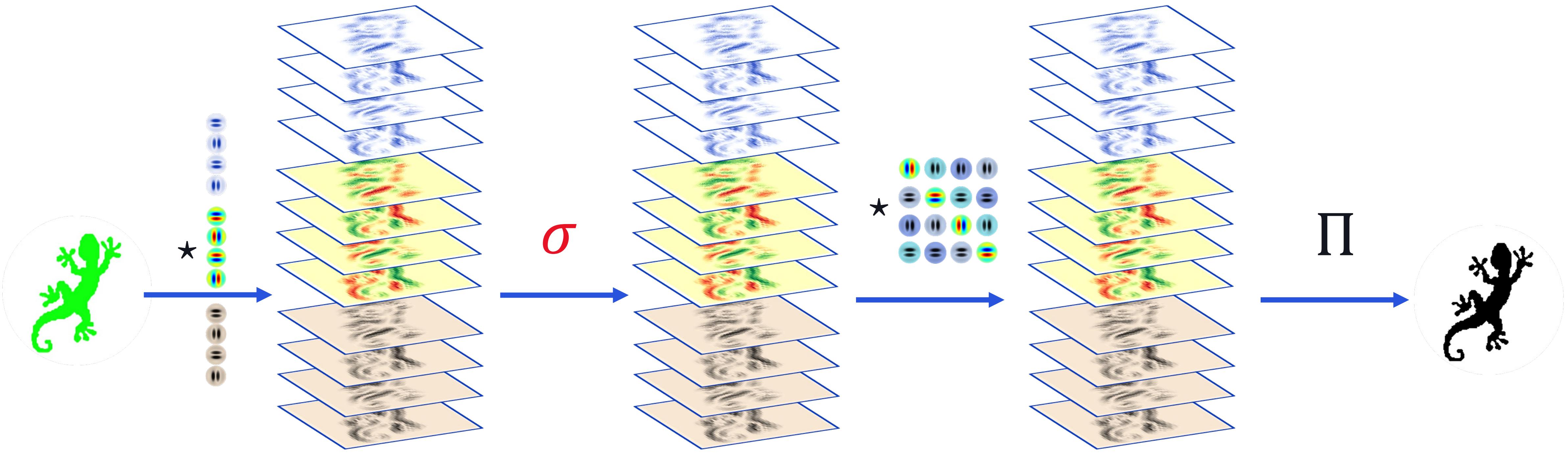
Is the output *rotation invariant*?

What is missing?



$f: \mathbb{Z}^n \to \mathbb{R}^{c_{in}}$ $\quad$ $f: G \to \mathbb{R}^c$ $\quad$ $f: G \to \mathbb{R}^c$ $\quad$ $f: \mathbb{Z}^n \to \mathbb{R}^{c_{out}}$

Activation Functions, Batch Normalization, …



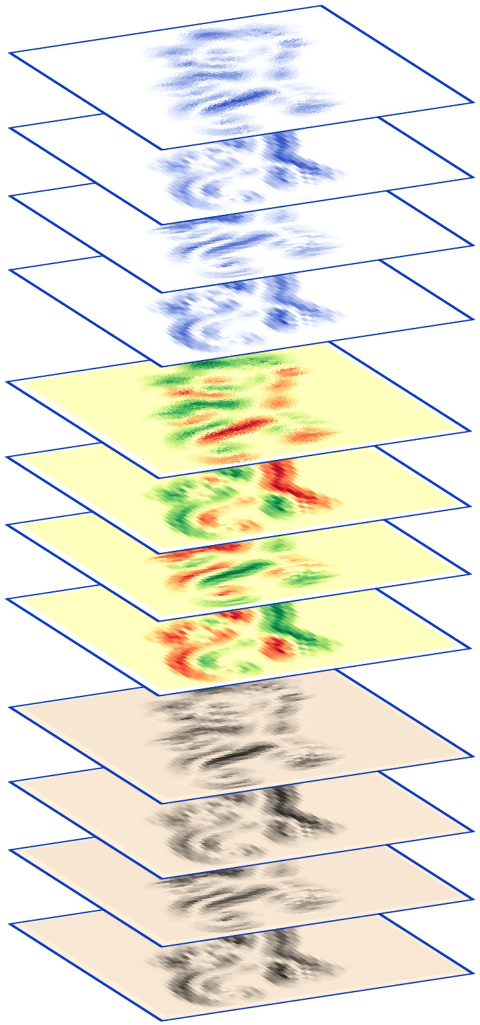
$f: \mathbb{Z}^n \to \mathbb{R}^{c_{in}}$ $f: G \to \mathbb{R}^c$ $f: G \to \mathbb{R}^c$ $f: G \to \mathbb{R}^c$ $f: \mathbb{Z}^n \to \mathbb{R}^{c_{out}}$

Activation Functions, Batch Normalization, …

Treat all ”group dimensions” as additional spatial dimensions

E.g. use `BatchNorm3D` instead of `BatchNorm2D` in `PyTorch`

Can use standard activations like `ReLU`



$$f : (\mathbb{Z}^2, +) \rtimes C_4 \to \mathbb{R}^3$$

GCNNs: Historical Notes

Convolution Neural Networks by Yann LeCun

GCNNs: Historical Notes

Convolution Neural Networks by Yann LeCun

Equivariant MLPs (Shawe-Taylor, 1989)

J. Shawe-Taylor. Building symmetries into feedforward networks. *IEE International Conference on Artificial Neural Networks, 1989.*

GCNNs: Historical Notes

Convolution Neural Networks by Yann LeCun

Equivariant MLPs (Shawe-Taylor, 1989)

GCNNs as introduced in the DL community by (Cohen and Welling, 2016)

Group Equivariant Convolutional Networks

Taco S. Cohen T.S.COHEN@UVA.NL
University of Amsterdam

Max Welling M.WELLING@UVA.NL
University of Amsterdam
University of California Irvine
Canadian Institute for Advanced Research

J. Shawe-Taylor. Building symmetries into feedforward networks. *IEE International Conference on Artificial Neural Networks,1989.*
Taco S. Cohen and Max Welling, Group Equivariant Convolutional Networks, *International Conference on Machine Learning (ICML), 2016*

GCNNs: Historical Notes

Convolution Neural Networks by Yann LeCun

Equivariant MLPs (Shawe-Taylor, 1989)

GCNNs as introduced in the DL community by (Cohen and Welling, 2016)

Similar ideas previously used in

- Scattering Convolution Networks by (Bruna and Mallat, 2013)

Invariant Scattering Convolution Networks

Joan Bruna and Stéphane Mallat
CMAP, Ecole Polytechnique, Palaiseau, France

J. Shawe-Taylor. Building symmetries into feedforward networks. *IEE International Conference on Artificial Neural Networks,1989.*
Joan Bruna and Stephane Mallat, Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2013.*
Taco S. Cohen and Max Welling, Group Equivariant Convolutional Networks, *International Conference on Machine Learning (ICML), 2016*

GCNNs: Historical Notes

Convolution Neural Networks by Yann LeCun

Equivariant MLPs (Shawe-Taylor, 1989)

GCNNs as introduced in the DL community by (Cohen and Welling, 2016)

Similar ideas previously used in

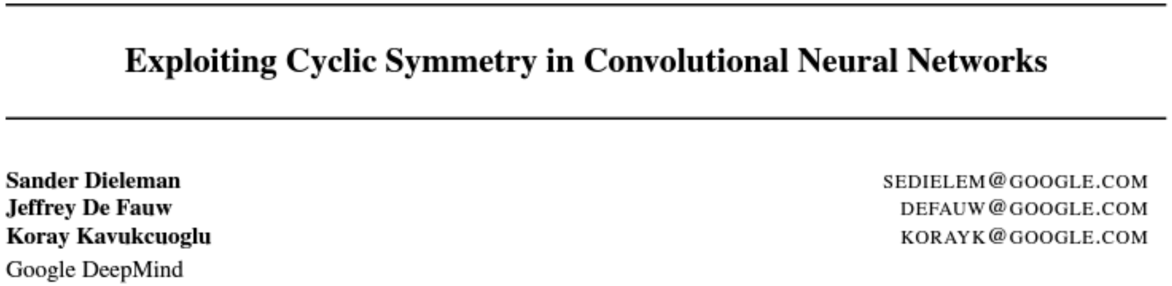
- Scattering Convolution Networks by (Bruna and Mallat, 2013)
- Deep Symmetry Networks by (Gens and Domingos, 2014)

Deep Symmetry Networks

Robert Gens **Pedro Domingos**
Department of Computer Science and Engineering
University of Washington
Seattle, WA 98195-2350, U.S.A.
{rcg,pedrod}@cs.washington.edu

J. Shawe-Taylor. Building symmetries into feedforward networks. *IEE International Conference on Artificial Neural Networks,1989.*
Joan Bruna and Stephane Mallat, Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2013.*
Robert Gens and Pedro M. Domingos, Deep symmetry networks. *In Advances in neural information processing systems (NIPS), 2014,*
Taco S. Cohen and Max Welling, Group Equivariant Convolutional Networks, *International Conference on Machine Learning (ICML), 2016*

GCNNs: Historical Notes


Exploiting Cyclic Symmetry in Convolutional Neural Networks

Sander Dieleman SEDIELEM@GOOGLE.COM
Jeffrey De Fauw DEFAUW@GOOGLE.COM
Koray Kavukcuoglu KORAYK@GOOGLE.COM
Google DeepMind

Convolution Neural Networks by Yann LeCun

Equivariant MLPs (Shawe-Taylor, 1989)

GCNNs as introduced in the DL community by (Cohen and Welling, 2016)

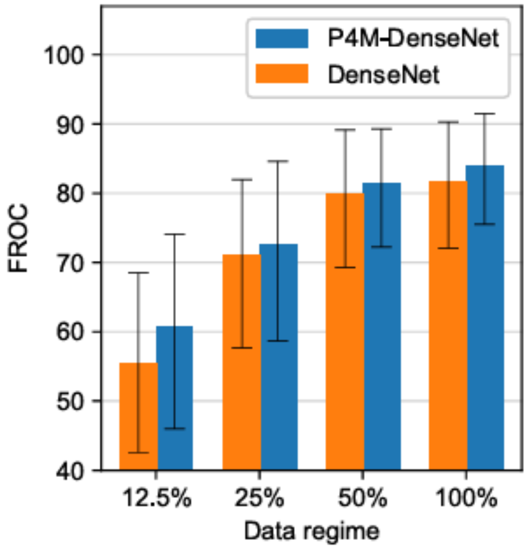
Similar ideas previously used in

- Scattering Convolution Networks by (Bruna and Mallat, 2013)
- Deep Symmetry Networks by (Gens and Domingos, 2014)
- $p4$ GCNN by (Dieleman, De Fauw and Kavukcuoglu, 2016)

J. Shawe-Taylor. Building symmetries into feedforward networks. *IEE International Conference on Artificial Neural Networks,1989.*
Joan Bruna and Stephane Mallat, Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2013*
Robert Gens and Pedro M. Domingos, Deep symmetry networks. *In Advances in neural information processing systems (NIPS), 2014*
Sander Dieleman, Jeffrey De Fauw and Koray Kavukcuoglu, Exploiting Cyclic Symmetry in Convolutional Neural Networks, *International Conference on Machine Learning (ICML), 2016*
Taco S. Cohen and Max Welling, Group Equivariant Convolutional Networks, *International Conference on Machine Learning (ICML), 2016*

GCNNs: Instantiations and Applications

Medical applications of $p4 = (\mathbb{Z}^2, +) \rtimes C_4$ and $p4m = (\mathbb{Z}^2, +) \rtimes D_4$ GCNNs



Performance on Camelyon16 dataset

Source: Rotation Equivariant CNNs for Digital Pathology, S. Veeling et al., 2018

Rotation Equivariant CNNs for Digital Pathology

Bastiaan S. Veeling*, Jasper Linmans*, Jim Winkens*, Taco Cohen, and Max Welling

University of Amsterdam, The Netherlands

Improved Semantic Segmentation for Histopathology using Rotation Equivariant Convolutional Networks

Jim Winkens *
University of Amsterdam
jim.winkens@gmail.com

Jasper Linmans *
University of Amsterdam
jasper.linmans@gmail.com

Bastiaan S. Veeling
University of Amsterdam
bas.veeling@gmail.com

Taco S. Cohen
University of Amsterdam
taco.cohen@gmail.com

Max Welling
University of Amsterdam
welling.max@gmail.com

GCNNs: Instantiations and Applications

Equivariance to isometries of other grids:

- HexaConv: $p6 = (\mathbb{Z}^2, +) \rtimes C_6$ and $p6m = (\mathbb{Z}^2, +) \rtimes D_6$ GCNNs

HEXACONV

Emiel Hoogeboom*, Jorn W.T. Peters* & Taco S. Cohen
University of Amsterdam
{e.hoogeboom,j.w.t.peters,t.s.cohen}@uva.nl

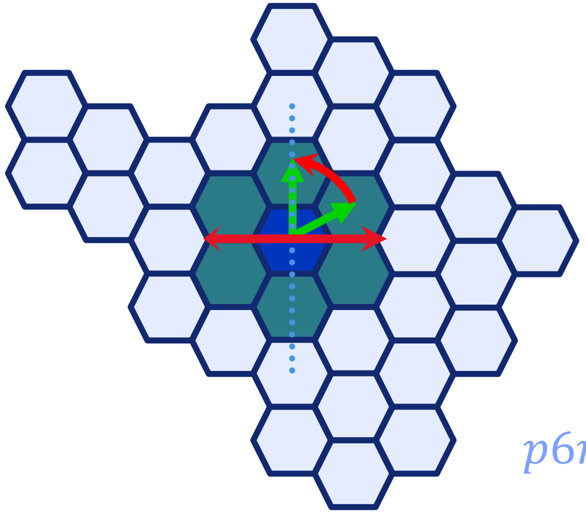
Max Welling
University of Amsterdam & CIFAR
m.welling@uva.nl

Table 1: CIFAR-10 performance comparison

Model	Error	Params
$\mathbb{Z}^2$	11.50 ±0.30	338000
$\mathbb{Z}^2$ Axial	11.25 ±0.24	337000
$p4$	10.08 ±0.23	337000
$p6$ Axial	9.98 ±0.32	336000
$p4m$	8.96 ±0.46	337000
$p6m$ Axial	8.64 ±0.34	337000

Table 2: AID performance comparison

Model	Error	Params
$\mathbb{Z}^2$	19.3 ±0.34	917000
$\mathbb{Z}^2$ Axial	17.8 ±0.37	916000
$p4$	10.7 ±0.36	915000
$p6$ Axial	8.7 ±0.72	916000
VGG (Transfer)	9.8 ±0.50	-

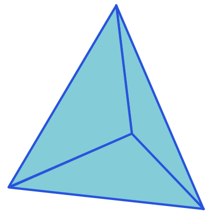


$p6m = (\mathbb{Z}^2, +) \rtimes D_6$

Natural images (Cifar10) and satellite images (Aerial Image Dataset, AID)

Hoogeboom, Emiel, Jorn WT Peters, Taco S. Cohen, and Max Welling. HexaConv. *In International Conference on Learning Representations (ICLR), 2018.*

GCNNs: Instantiations and Applications

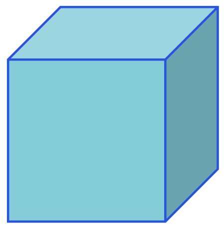


$T < O$ = 12 rotations of the tethrahedron

Equivariance to isometries of other grids:

- Cube symmetries
  - CubeNet $(\mathbb{Z}^3, +) \rtimes O$, $(\mathbb{Z}^3, +) \rtimes T$ and $(\mathbb{Z}^3, +) \rtimes D_2$
  - 3D G-CNNs $(\mathbb{Z}^3, +) \rtimes O$, $(\mathbb{Z}^3, +) \rtimes O_h$

    $(\mathbb{Z}^3, +) \rtimes D_4$ and $(\mathbb{Z}^3, +) \rtimes D_{4h}$



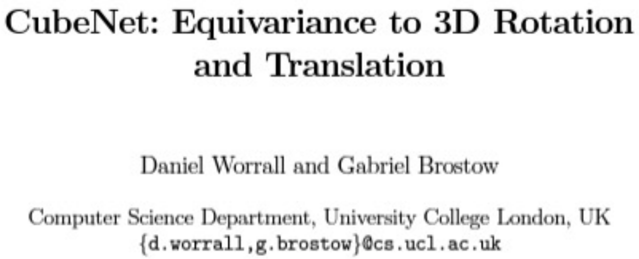
O = 24 rotations of the cube

Oh = 24 rotations + 24 reflections of the cube

$D_2 < O$ *Klein* group: subset of 4 rotations

$D_4 < O$ subset of 4 planar rotations and 1 out-of-plane rotation

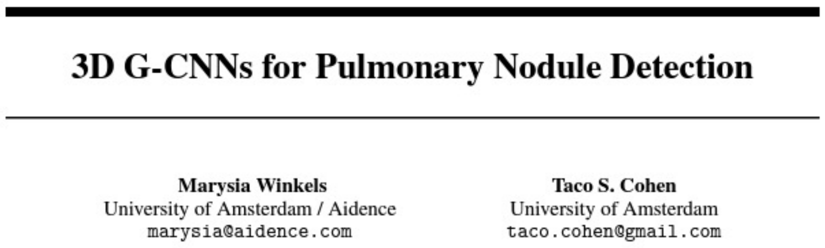
$D_{4h} = D_4 \rtimes C_2$ is D_4 combined with 3D inversions ($\cong C_2$)



CubeNet: Equivariance to 3D Rotation and Translation

Daniel Worrall and Gabriel Brostow

Computer Science Department, University College London, UK
{d.worrall,g.brostow}@cs.ucl.ac.uk



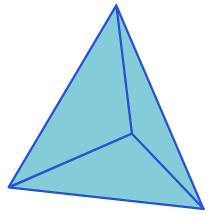
3D G-CNNs for Pulmonary Nodule Detection

Marysia Winkels
University of Amsterdam / Aidence
marysia@aidence.com

Taco S. Cohen
University of Amsterdam
taco.cohen@gmail.com

Daniel Worrall and Gabriel Brostow. CubeNet: Equivariance to 3D Rotation and Translation. *Proceedings of the European Conference on Computer Vision (ECCV), 2018*
Marysia Winkels and Taco S. Cohen. 3D G-CNNs for Pulmonary Nodule Detection. *Medical Imaging with Deep Learning, 2022*

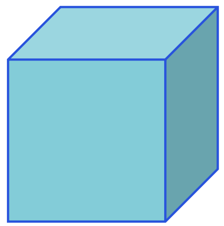
GCNNs: Instantiations and Applications



$T < O$ = 12 rotations of the tethrahedron

Equivariance to isometries of other grids:

- Cube symmetries
  - CubeNet $(\mathbb{Z}^3, +) \rtimes O$, $(\mathbb{Z}^3, +) \rtimes T$ and $(\mathbb{Z}^3, +) \rtimes D_2$
  - 3D G-CNNs $(\mathbb{Z}^3, +) \rtimes O$, $(\mathbb{Z}^3, +) \rtimes O_h$



N	$\mathbb{Z}^3$	D_4	D_{4h}	O	O_h
30	*0.252*	0.398	0.382	**0.562**	0.514
300	*0.550*	0.765	0.759	**0.767**	0.733
3,000	*0.791*	0.849	0.844	0.830	**0.850**
30,000	*0.843*	0.867	**0.880**	0.873	0.869

FROC score at different training set sizes N for different equivariance groups
Source: (Winkels and Cohen, 2022)

O = 24 rotations of the cube

Oh = 24 rotations + 24 reflections of the cube

$D_2 < O$ *Klein* group: subset of 4 rotations

$D_4 < O$ subset of 4 planar rotations and 1 out-of-plane rotation

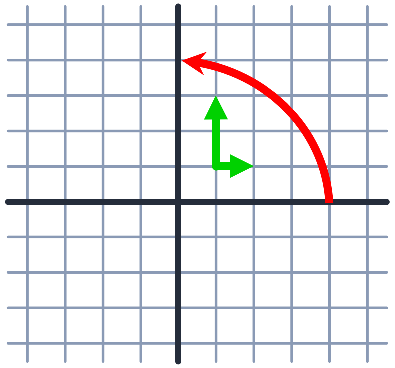
$D_{4h} = D_4 \rtimes C_2$ is D_4 combined with 3D inversions ($\cong C_2$)

Daniel Worrall and Gabriel Brostow. CubeNet: Equivariance to 3D Rotation and Translation. *Proceedings of the European Conference on Computer Vision (ECCV), 2018*
Marysia Winkels and Taco S. Cohen. 3D G-CNNs for Pulmonary Nodule Detection. *Medical Imaging with Deep Learning, 2022*

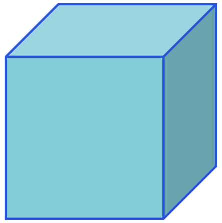
Question: did you notice any recurring pattern?

Question: did you notice any recurring pattern?

- ***Semi-direct*** product $G = (\mathbb{Z}^n, +) \rtimes H$
- Discrete rotation group H
- Euclidean grid $\mathbb{Z}^n$



$p4m = (\mathbb{Z}^2, +) \rtimes D_4$



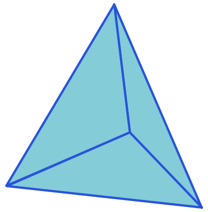
O = 24 rotations of the cube

Oh = 24 rotations + 24 reflections of the cube

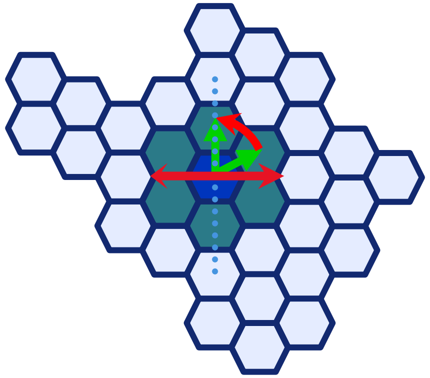
$D_2 < O$ *Klein* group: subset of 4 rotations

$D_4 < O$ subset of 4 planar rotations and 1 out-of-plane rotation

$D_{4h} = D_4 \rtimes C_2$ is D_4 combined with 3D inversions ($\cong C_2$)



$T < O$ = 12 rotations of the tethrahedron



$p6m = (\mathbb{Z}^2, +) \rtimes D_6$

$(\mathbb{Z}^3, +) \rtimes T$, $(\mathbb{Z}^3, +) \rtimes O$, $(\mathbb{Z}^3, +) \rtimes O_h$, $(\mathbb{Z}^3, +) \rtimes D_4$ and $(\mathbb{Z}^3, +) \rtimes D_{4h}$

Definition: **Semi-Direct product**

Given two groups G and H, and an action $\psi: H \times G \to G$ of H on G, the **semi-direct product** $G \rtimes_{\psi} H$ is a group with

- Elements: $G \times H = \{(g, h) | \, g \in G, h \in H\}$
- binary operation $\cdot: G \rtimes H \to G \rtimes H$

$$(g_2, h_2) \cdot (g_1, h_1) = (g_2 \, \psi(h_2, g_1), h_2 h_1)$$

Example: $SE(2) = (\mathbb{R}^2, +) \rtimes SO(2)$

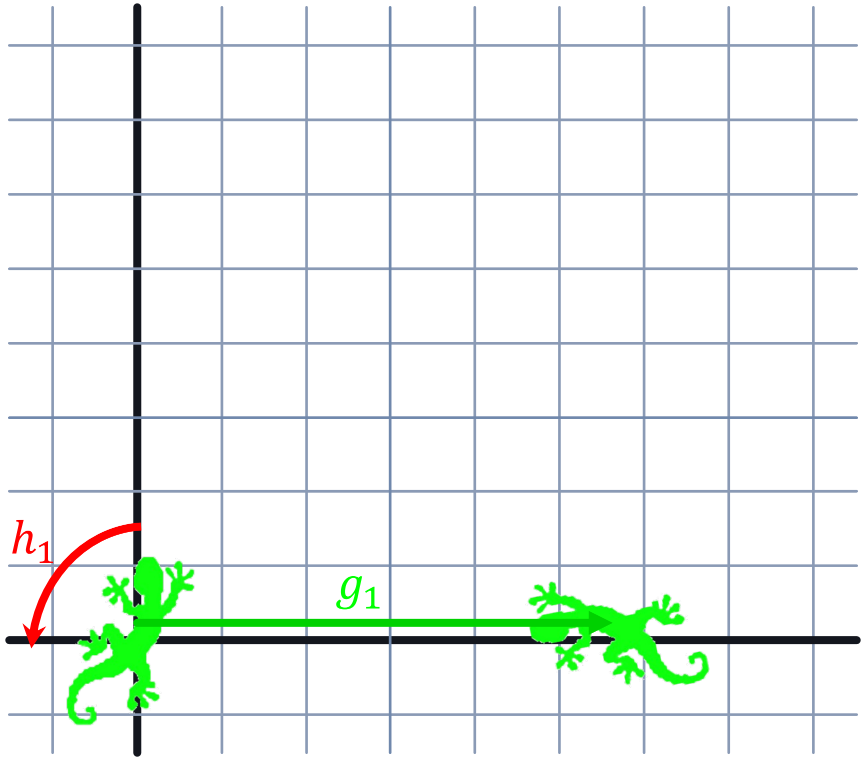
Definition: **Semi-Direct product**

Given two groups G and H, and an action $\psi: H \times G \to G$ of H on G, the **semi-direct product** $G \rtimes_\psi H$ is a group with

- Elements: $G \times H = \{(g, h) | \ g \in G, h \in H\}$
- binary operation $\cdot : G \rtimes H \to G \rtimes H$

$$(g_2, h_2) \cdot (g_1, h_1) = (g_2 \, \psi(h_2, g_1), h_2 h_1)$$

Example: $SE(2) = (\mathbb{R}^2, +) \rtimes SO(2)$



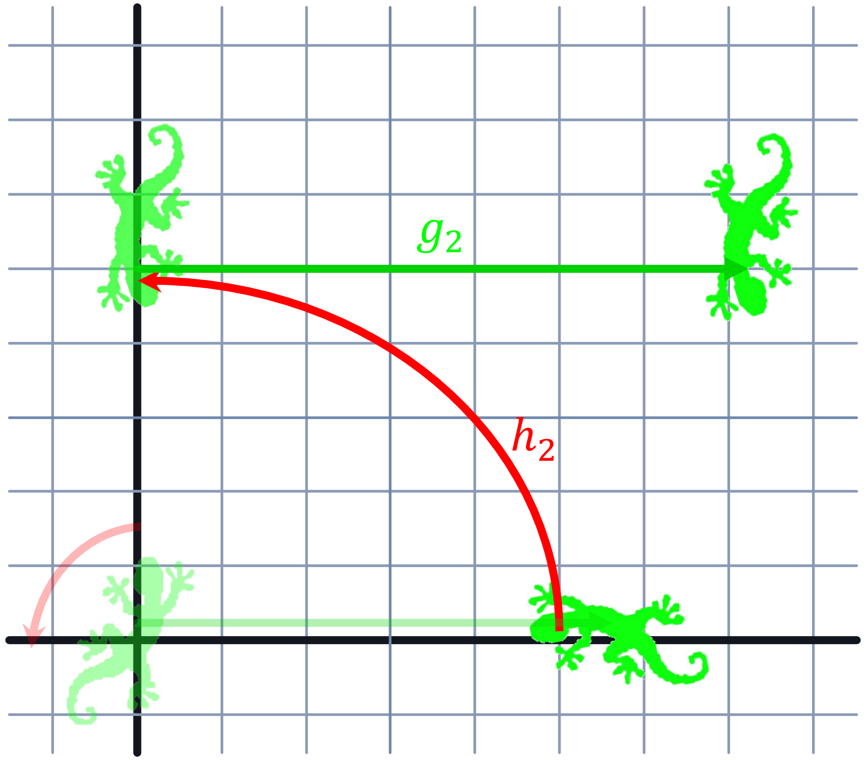
Definition: **Semi-Direct product**

Given two groups G and H, and an action $\psi: H \times G \to G$ of H on G, the **semi-direct product** $G \rtimes_{\psi} H$ is a group with

- Elements: $G \times H = \{(g, h) \mid g \in G, h \in H\}$
- binary operation $\cdot: G \rtimes H \to G \rtimes H$

$$(g_2, h_2) \cdot (g_1, h_1) = (g_2\, \psi(h_2, g_1), h_2 h_1)$$

Example: $SE(2) = (\mathbb{R}^2, +) \rtimes SO(2)$



Definition: **Semi-Direct product**

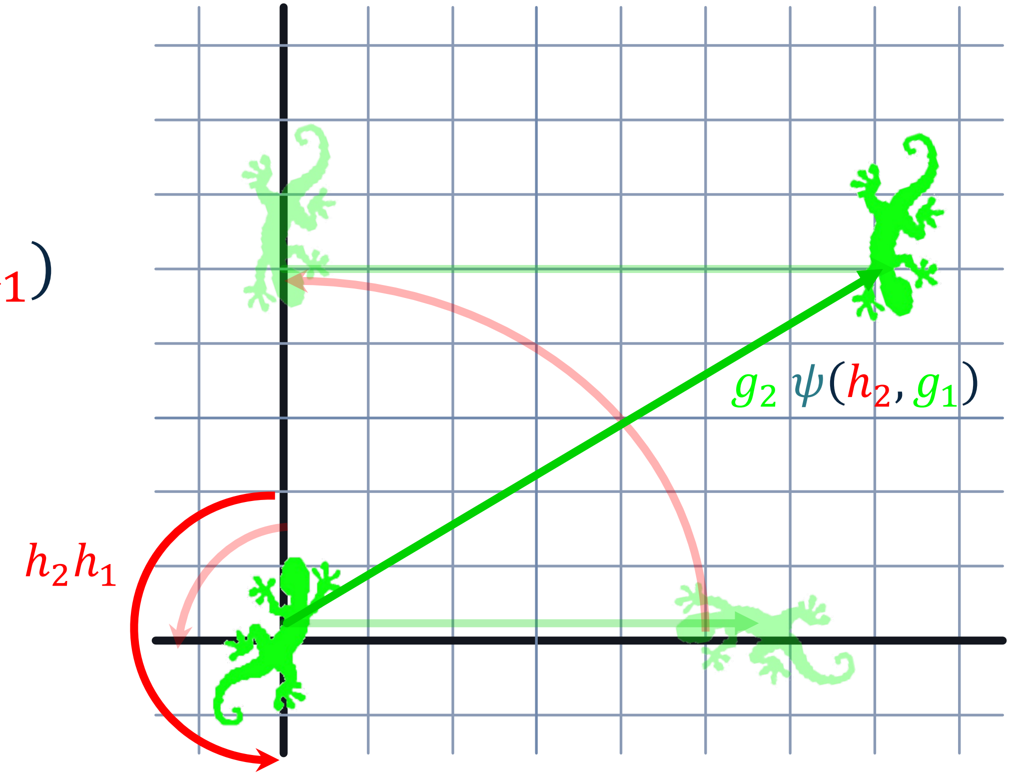
Given two groups G and H, and an action $\psi: H \times G \to G$ of H on G, the

semi-direct product $G \rtimes_\psi H$ is a group with

- Elements: $G \times H = \{(g, h) | \; g \in G, h \in H\}$
- binary operation $\cdot: G \rtimes H \to G \rtimes H$

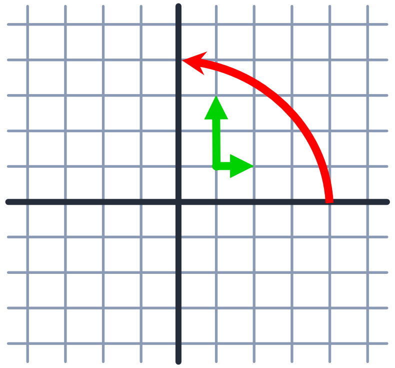
$$(g_2, h_2) \cdot (g_1, h_1) = (g_2 \, \psi(h_2, g_1), h_2 h_1)$$

Example: $SE(2) = (\mathbb{R}^2, +) \rtimes SO(2)$

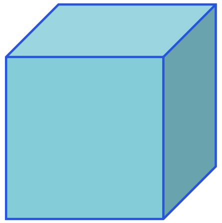


Question: did you notice any recurring pattern?

- ***Semi-direct*** product $G = (\mathbb{Z}^n, +) \rtimes H$
- Discrete rotation group H
- Euclidean grid $\mathbb{Z}^n$



$p4m = (\mathbb{Z}^2, +) \rtimes D_4$



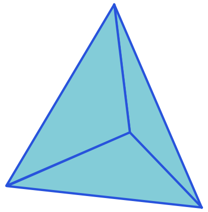
O = 24 rotations of the cube

Oh = 24 rotations + 24 reflections of the cube

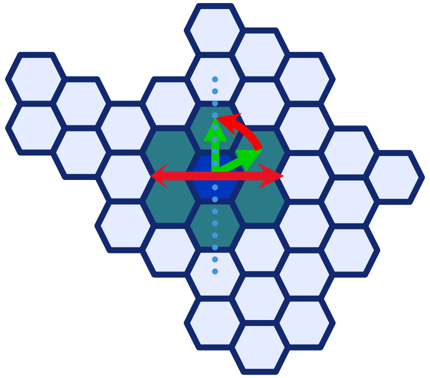
$D_2 < O$ *Klein* group: subset of 4 rotations

$D_4 < O$ subset of 4 planar rotations and 1 out-of-plane rotation

$D_{4h} = D_4 \rtimes C_2$ is D_4 combined with 3D inversions ($\cong C_2$)



$T < O$ = 12 rotations of the tethrahedron



$p6m = (\mathbb{Z}^2, +) \rtimes D_6$

$(\mathbb{Z}^3, +) \rtimes T, (\mathbb{Z}^3, +) \rtimes O, (\mathbb{Z}^3, +) \rtimes O_h, (\mathbb{Z}^3, +) \rtimes D_4$ and $(\mathbb{Z}^3, +) \rtimes D_{4h}$

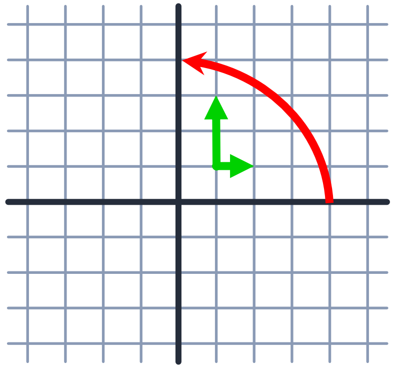
Group Convolution: Challenges

$$[\kappa \star f](g) = \int_{x \in X} \kappa(g^{-1}\,\boldsymbol{x})\; f(\boldsymbol{x}) d\mu(x)$$

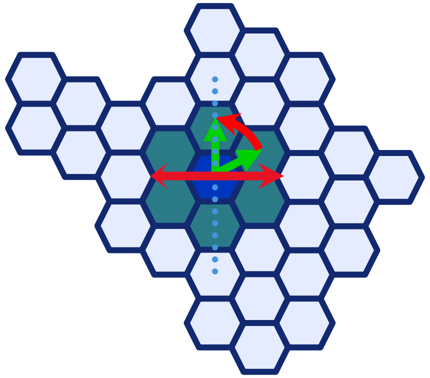
Integrate over X

Transform filter $\kappa \in L^2(X)$
by any group element
$g \in G$

Parameterize
features f
and filters $\kappa \in L^2(X)$



$p4m = (\mathbb{Z}^2, +) \rtimes D_4$



$p6m = (\mathbb{Z}^2, +) \rtimes D_6$

Semi-Direct Product group: *lifting convolution*

- ***Semi-direct*** product $G = (\mathbb{Z}^n, +) \rtimes H$

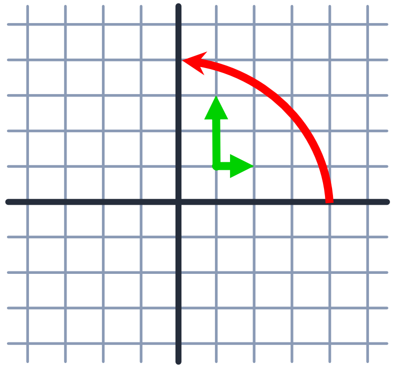
$$[\kappa \star f](t, h) = \sum_{x \in X = \mathbb{Z}^n} \kappa\left(h^{-1}(x - t)\right) f(x)$$

Sum over discrete grid

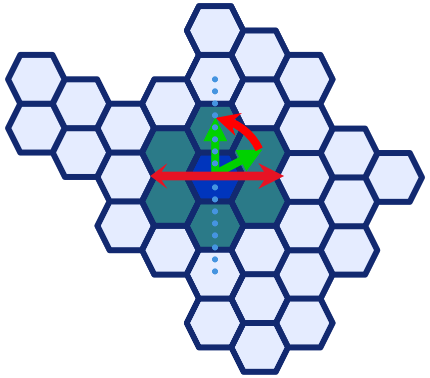
Normal convolution over $\mathbb{Z}^n$
with copies of filter κ
rotated only by H

Standard features f
discretized over
pixel/voxel grid

H only contains perfect symmetries of $\mathbb{Z}^n$



$p4m = (\mathbb{Z}^2, +) \rtimes D_4$

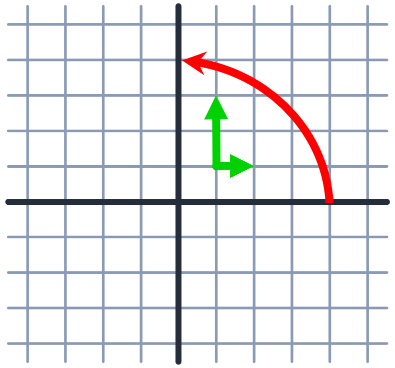


$p6m = (\mathbb{Z}^2, +) \rtimes D_6$

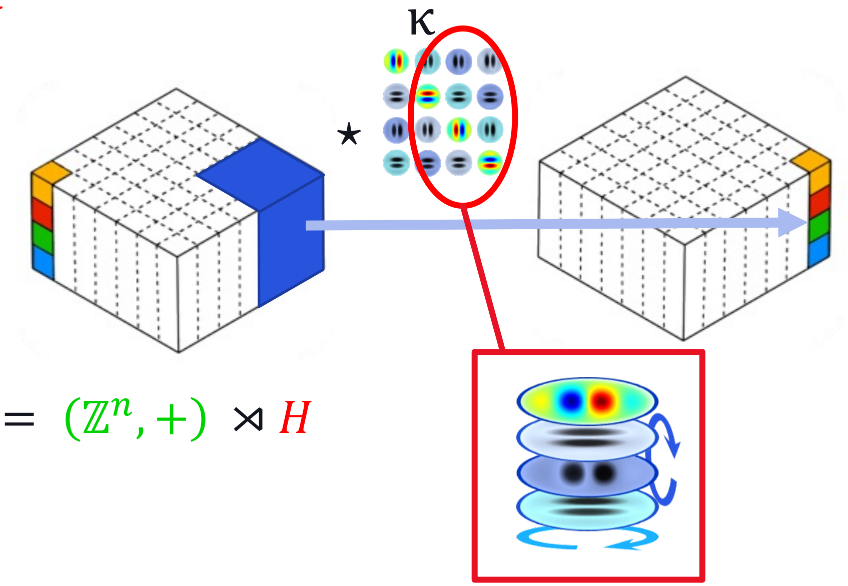
Semi-Direct Product group: *group convolution*

- ***Semi-direct*** product $G = (\mathbb{Z}^n, +) \rtimes H$

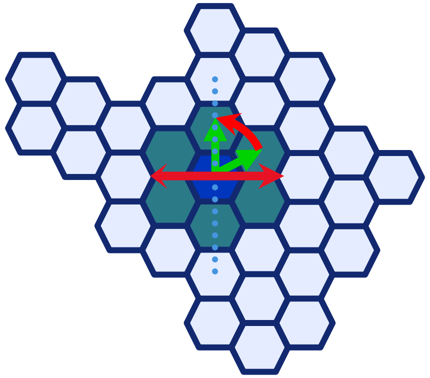
$$[\kappa \star f](t, h) = \sum_{\substack{x \in \mathbb{Z}^n \\ h' \in H}} \kappa(h^{-1}(x - t), h^{-1}h')\ f(x, h')$$



$p4m = (\mathbb{Z}^2, +) \rtimes D_4$



$X = G = (\mathbb{Z}^n, +) \rtimes H$



$p6m = (\mathbb{Z}^2, +) \rtimes D_6$

Finer rotation equivariance: beyond the symmetries of $\mathbb{Z}^n$

Discretize $SE(2) = (\mathbb{R}^2, +) \rtimes SO(2)$ rather than using $G = (\mathbb{Z}^n, +) \rtimes H$ convolution

$$[\kappa \star f](t, h) = \sum_{\boldsymbol{x} \in \mathbb{Z}^2} \kappa\left(h^{-1}(\boldsymbol{x} - t)\right) f(\boldsymbol{x})$$

Can use a larger discrete group
$H = C_n \subset SO(2)$

$\approx \int_{x \in \mathbb{R}^2} dx$

Rotate filters e.g. via interpolation

Finer rotation equivariance: beyond the symmetries of $\mathbb{Z}^n$

Discretize $SE(2) = (\mathbb{R}^2, +) \rtimes SO(2)$ rather than using $G = (\mathbb{Z}^n, +) \rtimes H$ convolution

- Learn one filter κ and rotate it via interpolation (Bekkers et al., 2018)
- Use an analytical basis to parameterize filters (Weiler et al., 2018)
  - Rotate analytical filters before sampling/discretizing them
  - Properly bandlimited basis limits discretization artifacts

Roto-Translation Covariant Convolutional Networks for Medical Image Analysis

Erik J Bekkers[*,1], Maxime W Lafarge[*,2], Mitko Veta[2], Koen AJ Eppenhof[2], Josien PW Pluim[2] and Remco Duits[1]

Eindhoven University of Technology, [1]Department of Mathematics and Computer Science and [2]Department of Biomedical Engineering, Eindhoven, The Netherlands
[*]Joint main authors - e.j.bekkers@tue.nl, m.w.lafarge@tue.nl

Learning Steerable Filters for Rotation Equivariant CNNs

Maurice Weiler[1,2] Fred A. Hamprecht[2] Martin Storath[2]
[1]AMLab / QUVA Lab, University of Amsterdam [2]HCI/IWR, University of Heidelberg
m.weiler@uva.nl {fred.hamprecht, martin.storath}@iwr.uni-heidelberg.de

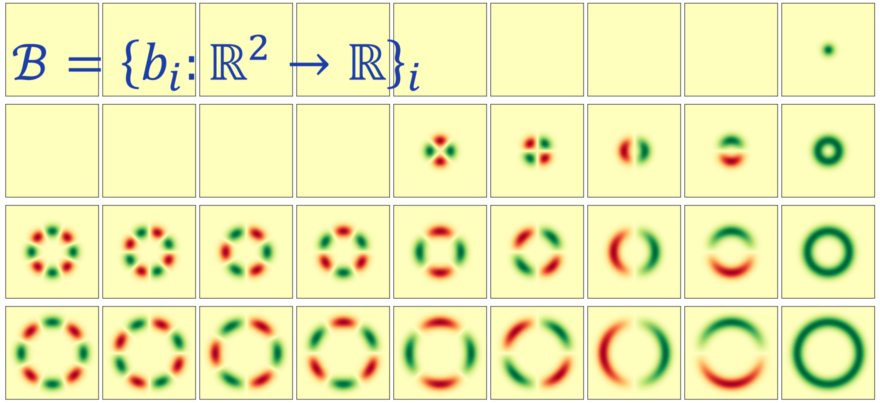
Erik J Bekkers, Maxime W Lafarge, Mitko Veta, Koen AJ Eppenhof, Josien PW Pluim, and Remco Duits. Roto-translation covariant convolutional networks for medical image analysis. *Medical Image Computing and Computer Assisted Intervention (MICCAI), 2018*
Maurice Weiler, Fred A. Hamprecht and Martin Storath. Learning steerable filters for rotation equivariant CNNs. *Conference on Computer Vision and Pattern Recognition (CVPR), 2018*

Finer rotation equivariance: beyond the symmetries of $\mathbb{Z}^n$

Discretize $SE(2) = (\mathbb{R}^2, +) \rtimes SO(2)$ rather than using $G = (\mathbb{Z}^n, +) \rtimes H$ convolution

- Learn one filter κ and rotate it via interpolation (Bekkers et al., 2018)
- Use an analytical basis to parameterize filters (Weiler et al., 2018)
  - Rotate analytical filters before sampling/discretizing them
  - Properly bandlimited basis limits discretization artifacts

$$[h.\kappa](\boldsymbol{x}) = \kappa(h^{-1}\boldsymbol{x}) = \sum_i w_i b_i(h^{-1}\boldsymbol{x})$$



Erik J Bekkers, Maxime W Lafarge, Mitko Veta, Koen AJ Eppenhof, Josien PW Pluim, and Remco Duits. Roto-translation covariant convolutional networks for medical image analysis. *Medical Image Computing and Computer Assisted Intervention (MICCAI), 2018*

Maurice Weiler, Fred A. Hamprecht and Martin Storath. Learning steerable filters for rotation equivariant CNNs. *Conference on Computer Vision and Pattern Recognition (CVPR), 2018*

Beyond the symmetries of $\mathbb{Z}^n$

Monte-Carlo approximation of the integral rather than using standard convolution

$$[\kappa \star f](g) = \frac{1}{|\mathcal{X}|} \sum_{\boldsymbol{x} \in \mathcal{X}} \kappa(g^{-1}x)\ f(\boldsymbol{x})$$

Samples from a finite subset
$\mathcal{G} \subset G$

$\approx \int_{x \in X} dx$

Samples from a finite subset
$\mathcal{X} \subset X$
(e.g. $\mathcal{X} = \mathcal{G}$ in intermidiate layers)

How can we parameterize filters?

Marc Finzi, Samuel Stanton, Pavel Izmailov, and Andrew Gordon Wilson. Generalizing convolutional neural networks for equivariance to Lie groups on arbitrary continuous data. *International Conference on Machine Learning (ICML), 2020.*

Beyond the symmetries of $\mathbb{Z}^n$

Monte-Carlo approximation of the integral rather than using standard convolution

$$[\kappa \star f](g) = \frac{1}{|\mathcal{X}|} \sum_{\boldsymbol{x} \in \mathcal{X}} \kappa(g^{-1}x)\ f(\boldsymbol{x})$$

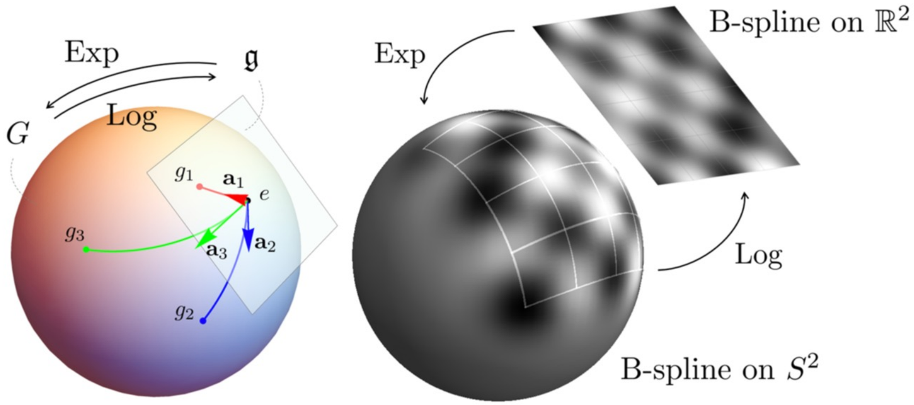


Image source: courtesy of Erik J Bekkers

Filter parameterized on **tangent space** of X

$$\kappa(x) = \kappa(\log x)$$

e.g. the Lie algebra $\mathfrak{g}$ of the Lie group G

- MLP (Finzi et al., 2020)
- B-Spline (Bekkers, 2020)

Marc Finzi, Samuel Stanton, Pavel Izmailov, and Andrew Gordon Wilson. Generalizing convolutional neural networks for equivariance to Lie groups on arbitrary continuous data. *International Conference on Machine Learning (ICML), 2020.*

Erik J Bekkers. B-spline CNNs on Lie groups. *International Conference on Learning Representations (ICLR), 2020.*

Beyond the symmetries of $\mathbb{Z}^n$

Monte-Carlo approximation of the integral rather than using standard convolution

$$[\kappa \star f](g) = \underbrace{\frac{1}{|\mathcal{X}|}}_{\approx \int_{x \in X} dx} \underbrace{\sum_{\boldsymbol{x} \in \mathcal{X}}}_{\text{Samples from a finite subset } \mathcal{X} \subset X} \kappa(g^{-1}x)\ f(\boldsymbol{x})$$

Equivariance “in expectation”

Discretization artifacts hard to control (especially with MLP filters)

Marc Finzi, Samuel Stanton, Pavel Izmailov, and Andrew Gordon Wilson. Generalizing convolutional neural networks for equivariance to Lie groups on arbitrary continuous data. *International Conference on Machine Learning (ICML), 2020.*

Erik J Bekkers. B-spline CNNs on Lie groups. *International Conference on Learning Representations (ICLR), 2020.*

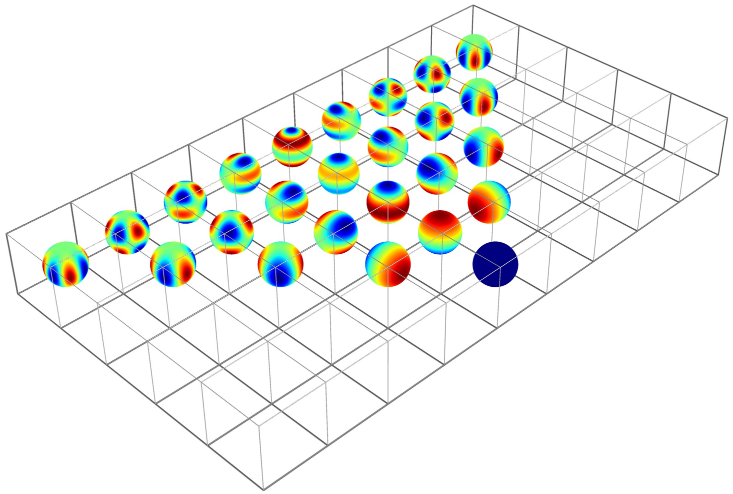
Beyond the symmetries of $\mathbb{Z}^n$: Spherical CNNs (Cohen et al., 2018)

Spherical signals in $L^2(S^2)$ and $G = SO(3)$ equivariance

Leverage *Fourier Transform* to

- Parameterize signals over $X = S^2$ and $X = SO(3)$
  - Filters and features defined by weights which linearly combine analytical (Fourier) basis for $L^2(X)$
- Efficiently perform convolution (*Fast Fourier Transform* and *Convolution Theorem*)

$$[h.\kappa](\boldsymbol{x}) = \kappa(h^{-1}\boldsymbol{x}) = \sum_i w_i b_i(h^{-1}\boldsymbol{x})$$



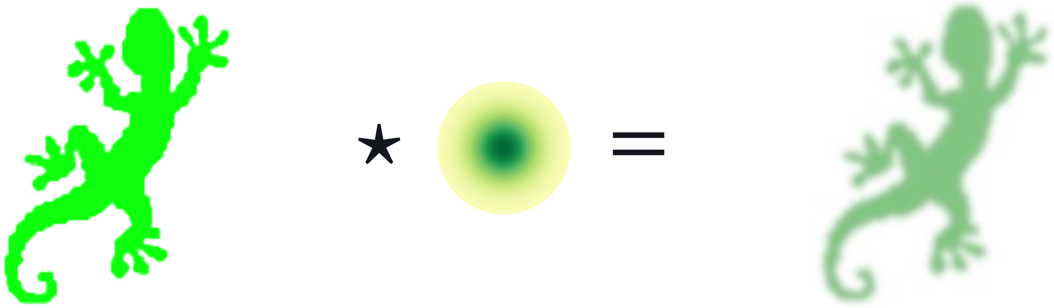
Discretization via Inverse FT to sample features and apply activation function σ

We will hear more about Fourier Transform over groups in the next lecture

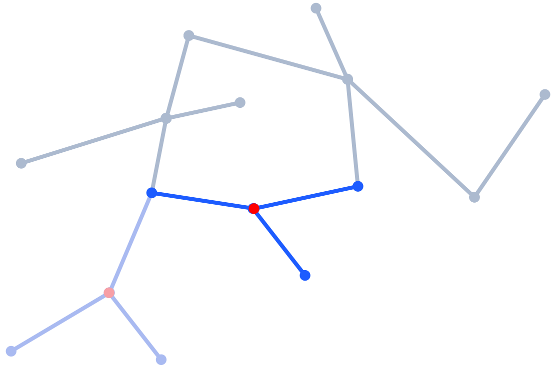
Taco S. Cohen, Mario Geiger, Jonas Köhler, and Max Welling. Spherical CNNs. *International Conference on Learning Representations (ICLR), 2018.*

Isotropic Filters

- Isotropic filters: the response does not change when rotated



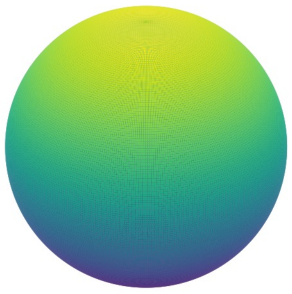
- Not very expressive
- Analogous to Graph Message Passing: no directional dependence!



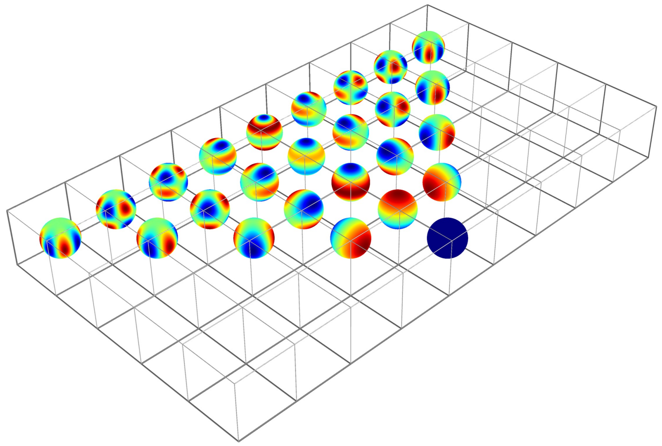
Spherical CNNs with Isotropic Filters (Esteves et al., 2018)

Isotropic filters parameterize convolution $L^2(S^2) \rightarrow L^2(S^2)$

Rather than convolution $L^2(S^2) \rightarrow L^2\big(SO(3)\big)$



vs



Trades off expressivity for efficiency

Carlos Esteves, Christine Allen-Blanchette, Ameesh Makadia, and Kostas Daniilidis. Learning SO(3) equivariant representations with spherical CNNs. *European Conference on Computer Vision (ECCV), 2018.*

Is Group Convolution all we need?

We have seen that Group Convolution provides an **equivariant linear map**

$$[\kappa \star \cdot]: L^2(X) \to L^2(G)$$

Are there any other such maps?

Group Convolution is all you need

We have seen that Group Convolution provides an **equivariant linear map**

$$[\kappa \star \cdot]: L^2(X) \to L^2(G)$$

Group-CNNs are *universal approximators* for equivariant functions

Siamak Ravanbakhsh. Universal Equivariant Multilayer Perceptrons. *International Conference on Machine Learning (ICML), 2020.*
Haggai Maron, Ethan Fetaya, Nimrod Segol, and Yaron Lipman. On the Universality of Invariant Networks. *International Conference on Machine Learning (ICML), 2019*
Wataru Kumagai and Akiyoshi Sannai. Universal Approximation Theorem for Equivariant Maps by Group CNNs, *2020*
Sho Sonoda, Isao Ishikawa, and Masahiro Ikeda. Universality of group convolutional neural networks based on ridgelet analysis on groups. *Advances in Neural Information Processing Systems (NeurIPS), 2022*

Group Convolution is all you need

We have seen that Group Convolution provides an **equivariant linear map**

$$[\kappa \star \cdot]: L^2(X) \to L^2(G)$$

Any equivariant linear map of this form is a convolution! (typically, under some minor conditions)

- For any **compact** group G (Kondor and Trivedi, 2018), (Cohen et al., 20120)
- For any **Lie** group (Bekker, 2020)
- Unimodular Lie groups (Aronsson, 2022)

Risi Kondor and Shubhendu Trivedi. On the generalization of equivariance and convolution in neural networks to the action of compact groups. *International Conference on Machine Learning (ICML), 2018.*

Taco S. Cohen, Mario Geiger, and Maurice Weiler. A general theory of equivariant CNNs on homogeneous spaces. *Advances in Neural Information Processing Systems (NeurIPS), 2019*

Erik J Bekkers. B-spline CNNs on Lie groups. *International Conference on Learning Representations (ICLR), 2020.*

Jimmy Aronsson. Homogeneous vector bundles and G-equivariant convolutional neural networks. *Sampling Theory, Signal Processing, and Data Analysis, 2022*

Thank you

Nothing in these materials is an offer to sell any of the components or devices referenced herein.

References in this presentation to “Qualcomm” may mean Qualcomm Incorporated, Qualcomm Technologies, Inc., and/or other subsidiaries or business units within the Qualcomm corporate structure, as applicable. Qualcomm Incorporated includes our licensing business, QTL, and the vast majority of our patent portfolio. Qualcomm Technologies, Inc., a subsidiary of Qualcomm Incorporated, operates, along with its subsidiaries, substantially all of our engineering, research and development functions, and substantially all of our products and services businesses, including our QCT semiconductor business.

Snapdragon and Qualcomm branded products are products of Qualcomm Technologies, Inc. and/or its subsidiaries.
Qualcomm patented technologies are licensed by Qualcomm Incorporated.

Follow us on:

For more information, visit us at qualcomm.com & qualcomm.com/blog