Group Equivariant Deep Learning

Lecture 3 - Equivariant graph neural networks

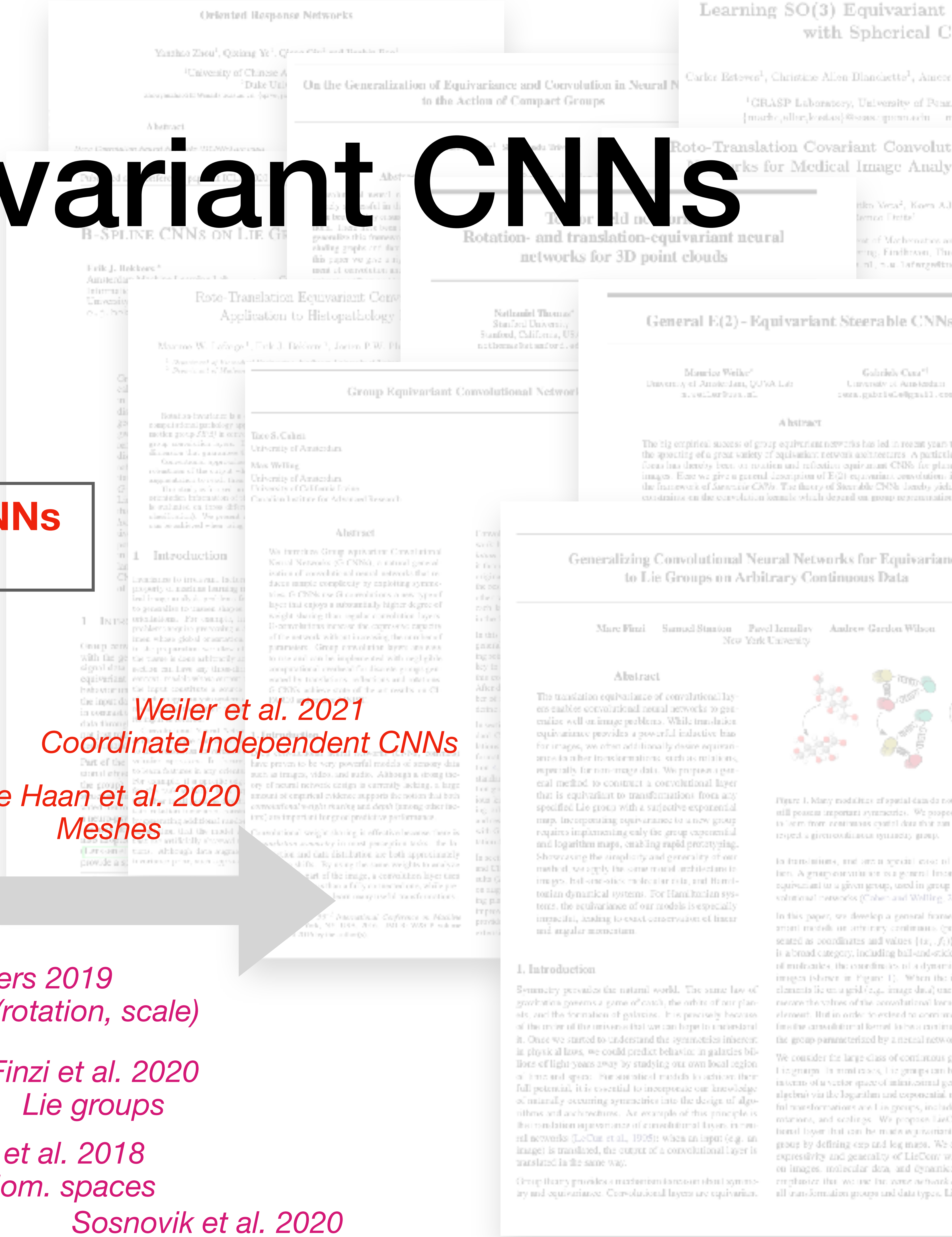
Lecture 3.7 - Gauge equivariant graph NNs

Erik Bekkers, Amsterdam Machine Learning Lab, University of Amsterdam

This mini-course serves as a module with the UvA Master AI course Deep Learning 2 https://uvadl2c.github.io/

From plain NNs to Gauge-equivariant CNNs

https://github.com/Chen-Cai-OSU/awesome-equivariant-network



Discrete G-CNNs
(square/cube symmetries)

Gauge-equivariant CNNs
(manifolds)

Cohen-Welling 2016
p4m

Winkels-Cohen 2018
3D discrete roto-transl.

Hoogeboom et al. 2018
SE(2,6)

Dieleman et al. 2016
p4m

Worrall-Brostow 2018
3D discrete roto-transl.

Cohen et al. 2019
Icosahedron

Weiler et al. 2021
Coordinate Independent CNNs

de Haan et al. 2020
Meshes

NNs → CNNs → G-CNNs

LeCun et al. 1990

Mallat et al. 2013, 2015
SE(2)

Bekkers et al. 2014-2018
SE(2)

Weiler et al. 2017
SE(2)

Zhou et al. 2017
SE(2)

Worrall et al. 2017
SE(2)

Esteves 2017
SO(3)

Kondor 2018
SE(3)

Cohen et al. 2018
SE(3)

Kondor-Trivedi 2018
SE(3)

Weiler et al. 2018
SE(3)

Thomas et al. 2018
SE(3)

Weiler-Cesa 2019
SE(2)

Bekkers 2019
Lie groups (rotation, scale)

Finzi et al. 2020
Lie groups

Chakraborty et al. 2018
Riemannian Hom. spaces

Sosnovik et al. 2020
Scale-translation

Continuous rotation G-CNNs
(regular discretizations)

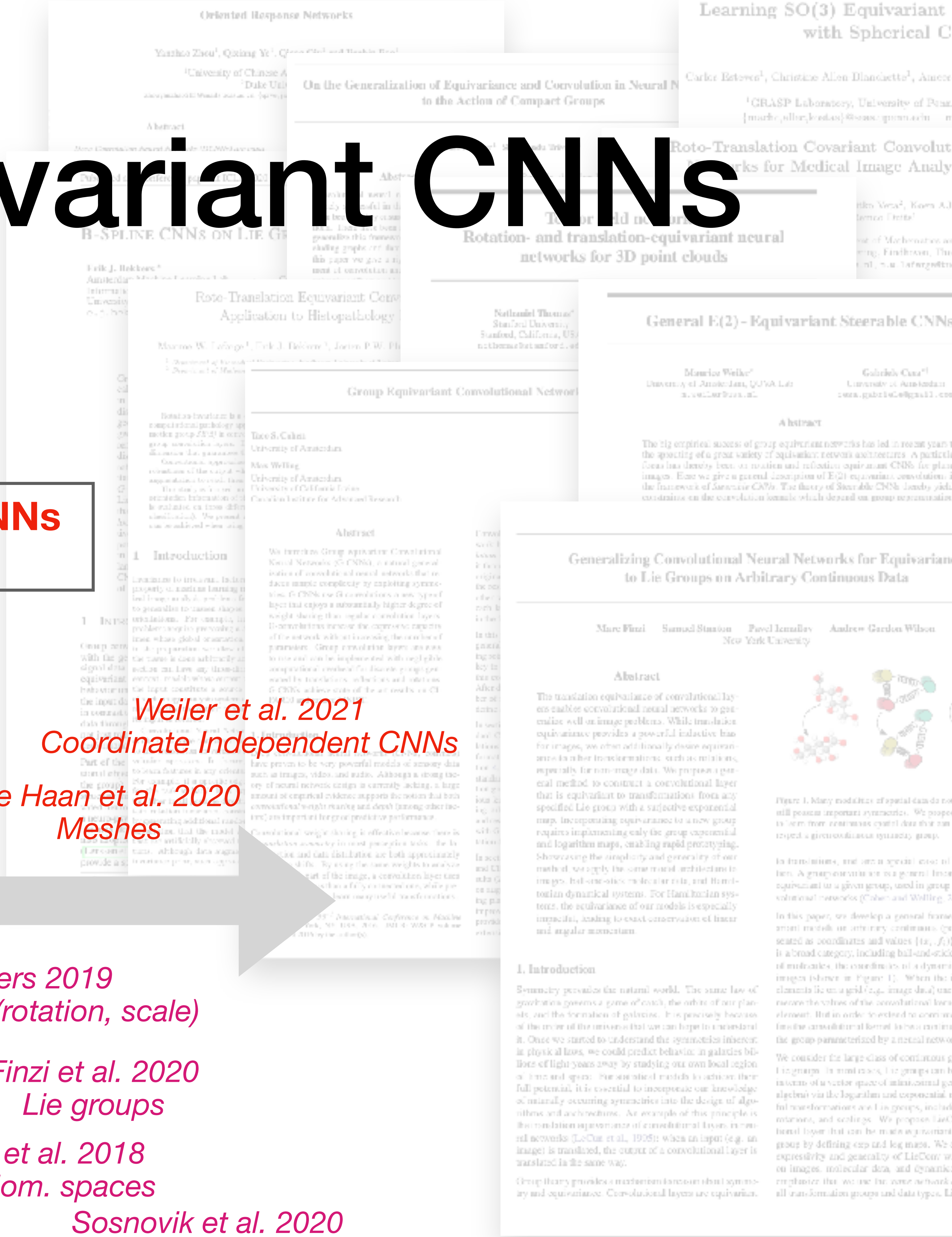
Continuous rotation G-CNNs
(steerable)

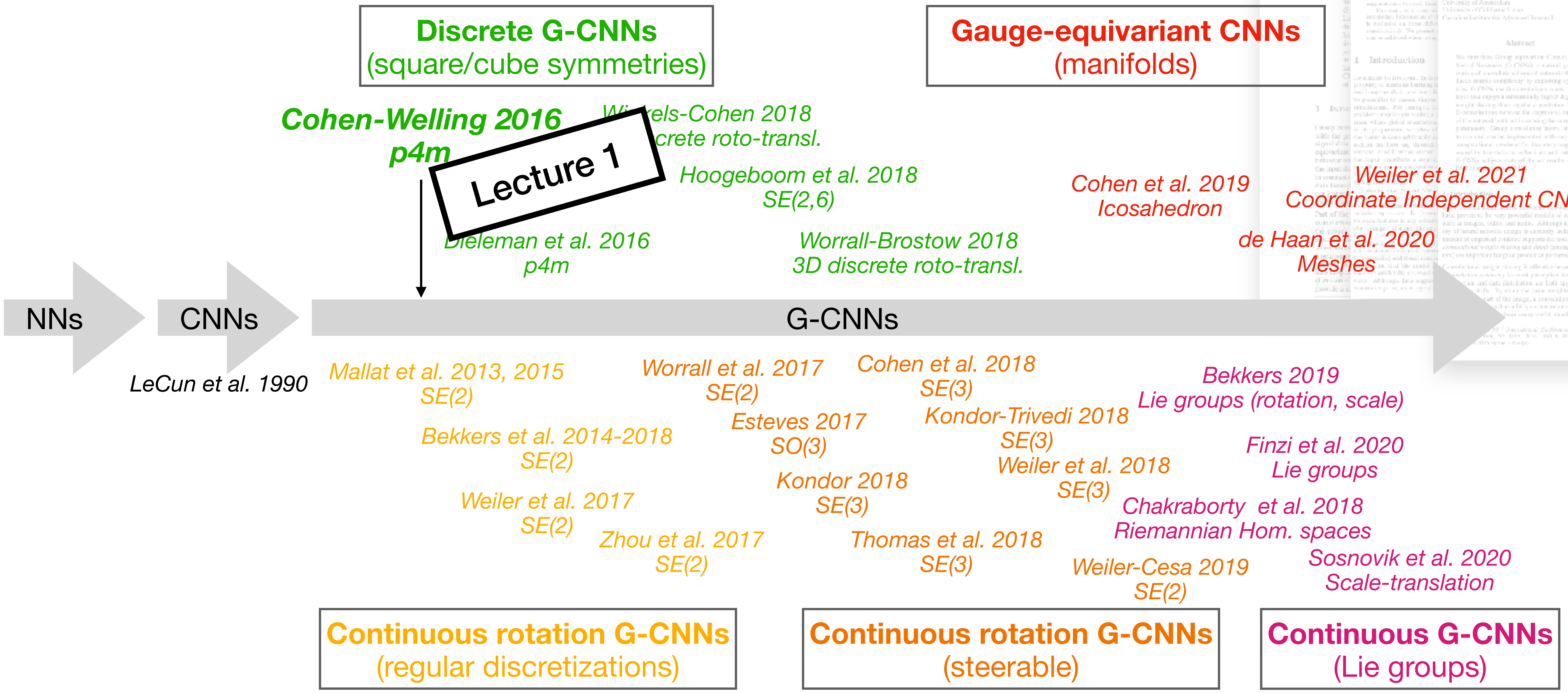
Continuous G-CNNs
(Lie groups)

Cesa-Lang-Weiler 2022
$G = \mathbb{R}^d \rtimes H$ *with* H *compact*

From plain NNs to Gauge-equivariant CNNs

https://github.com/Chen-Cai-OSU/awesome-equivariant-network

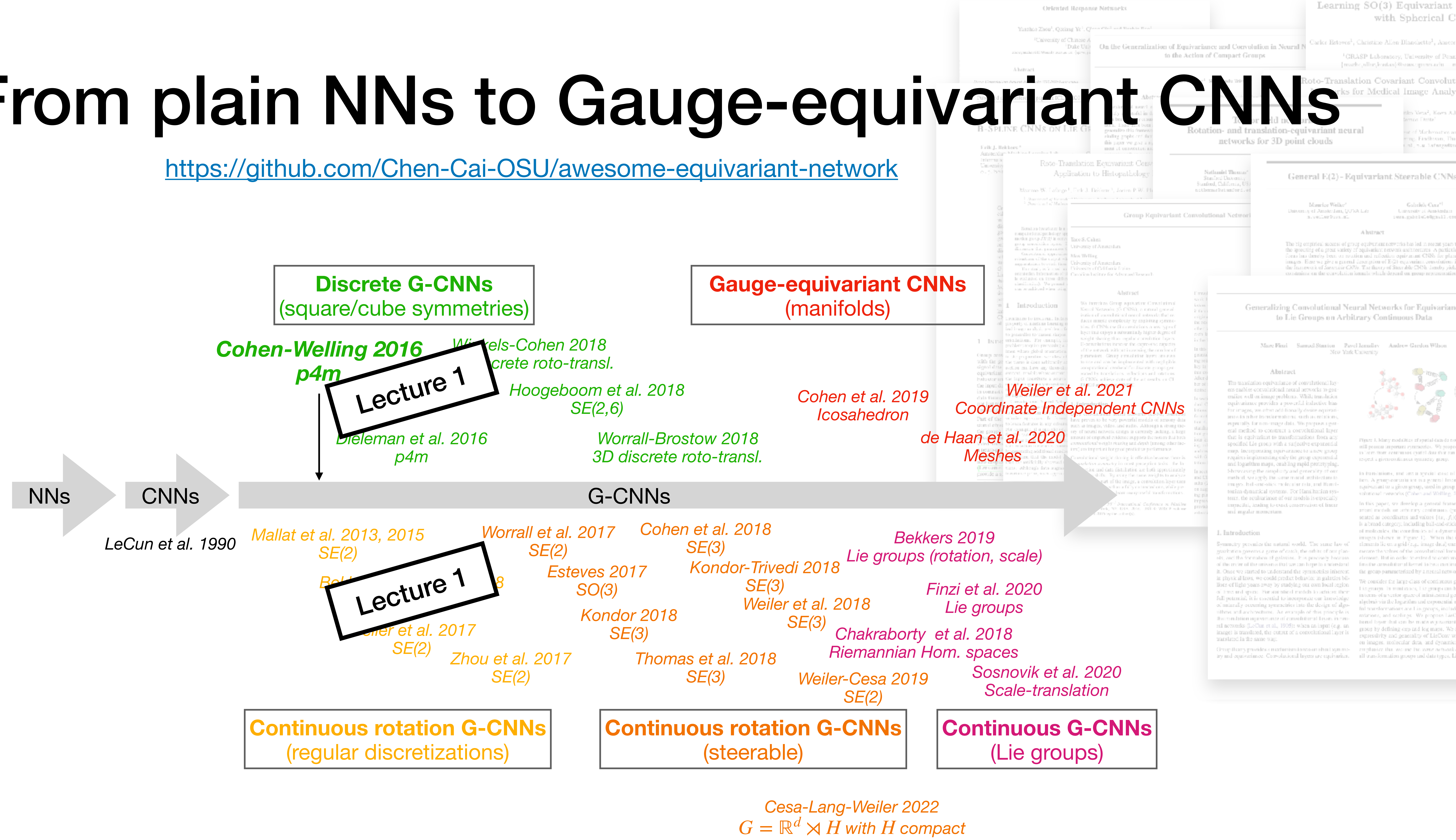




$G = \mathbb{R}^d \rtimes H$ *with* H *compact*

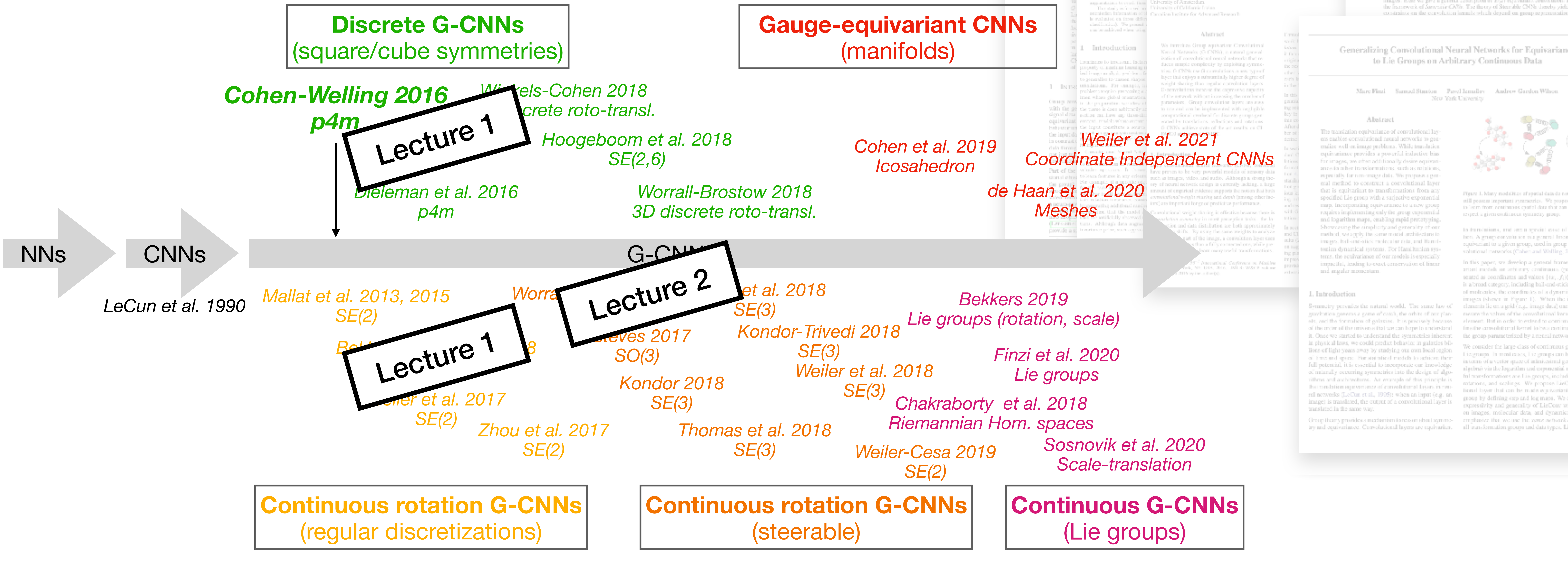
From plain NNs to Gauge-equivariant CNNs

https://github.com/Chen-Cai-OSU/awesome-equivariant-network



From plain NNs to Gauge-equivariant CNNs

https://github.com/Chen-Cai-OSU/awesome-equivariant-network

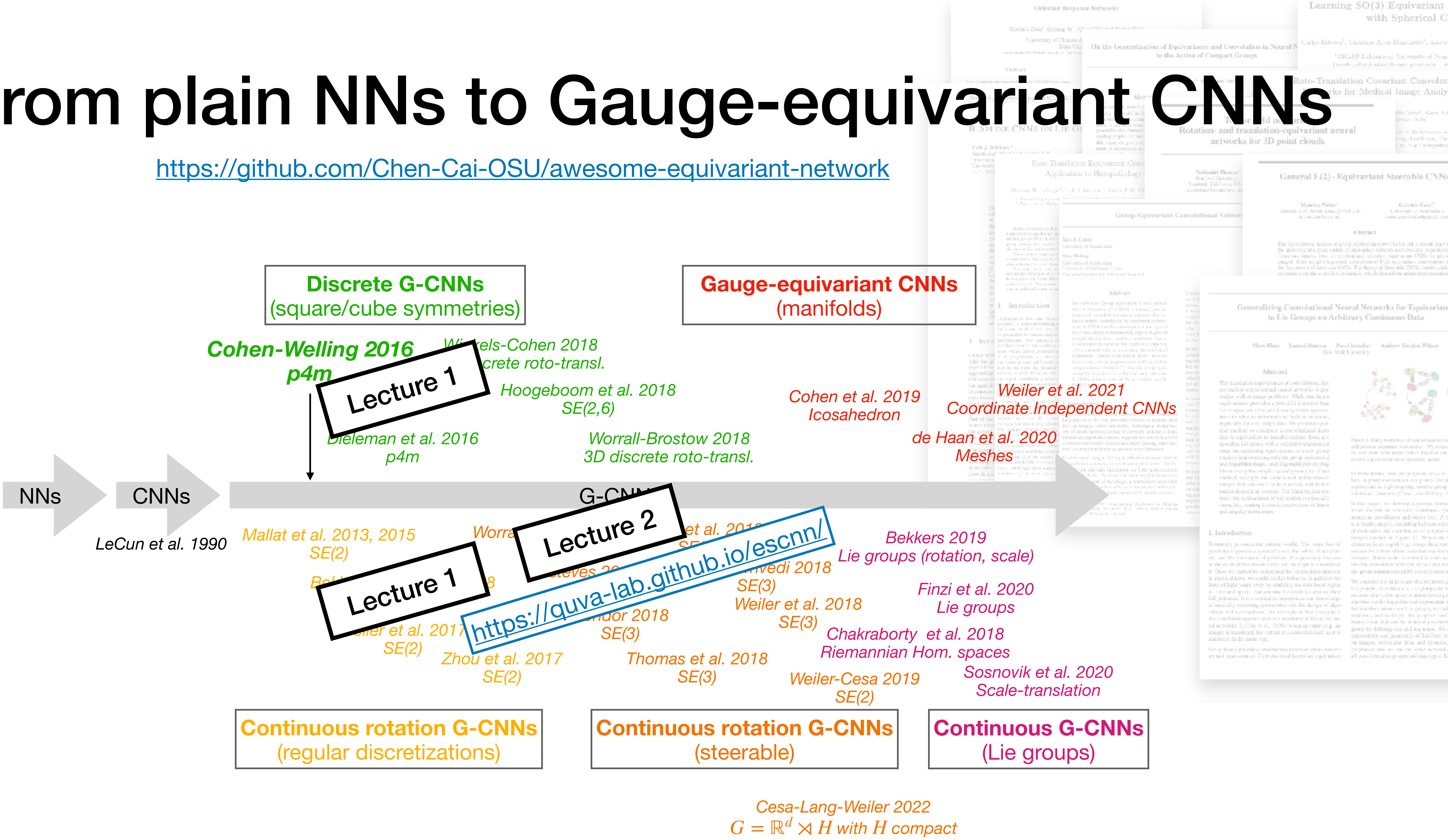


Cesa-Lang-Weiler 2022

$G = \mathbb{R}^d \rtimes H$ with H compact

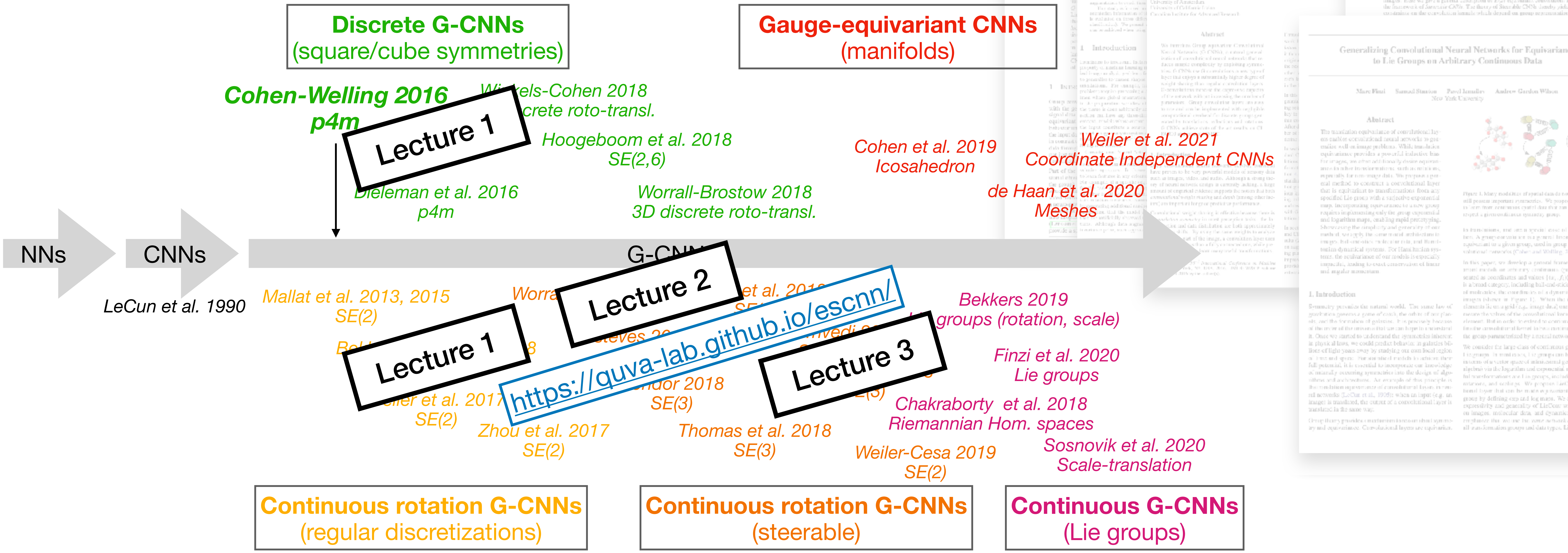
From plain NNs to Gauge-equivariant CNNs

https://github.com/Chen-Cai-OSU/awesome-equivariant-network



From plain NNs to Gauge-equivariant CNNs

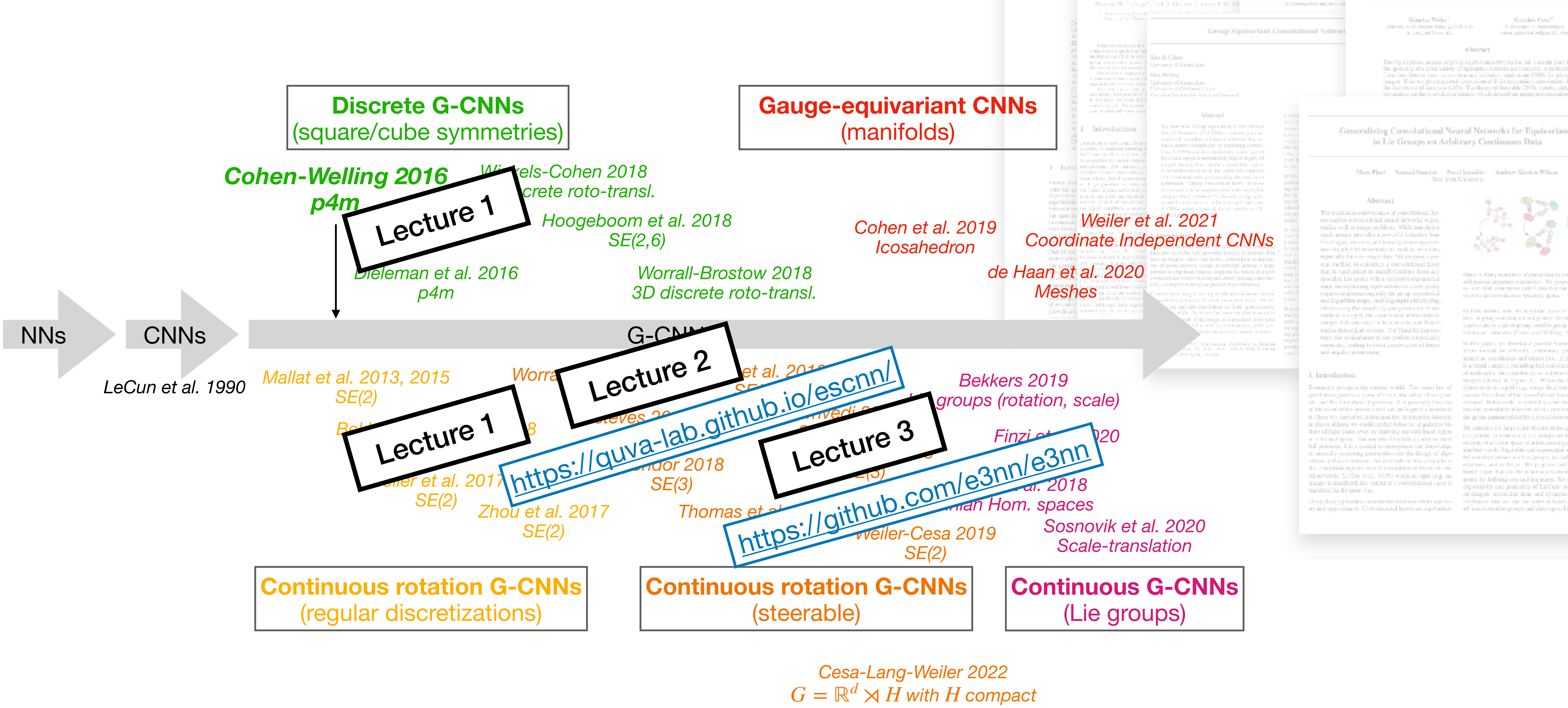
https://github.com/Chen-Cai-OSU/awesome-equivariant-network



Cesa-Lang-Weiler 2022
$G = \mathbb{R}^d \rtimes H$ with H compact

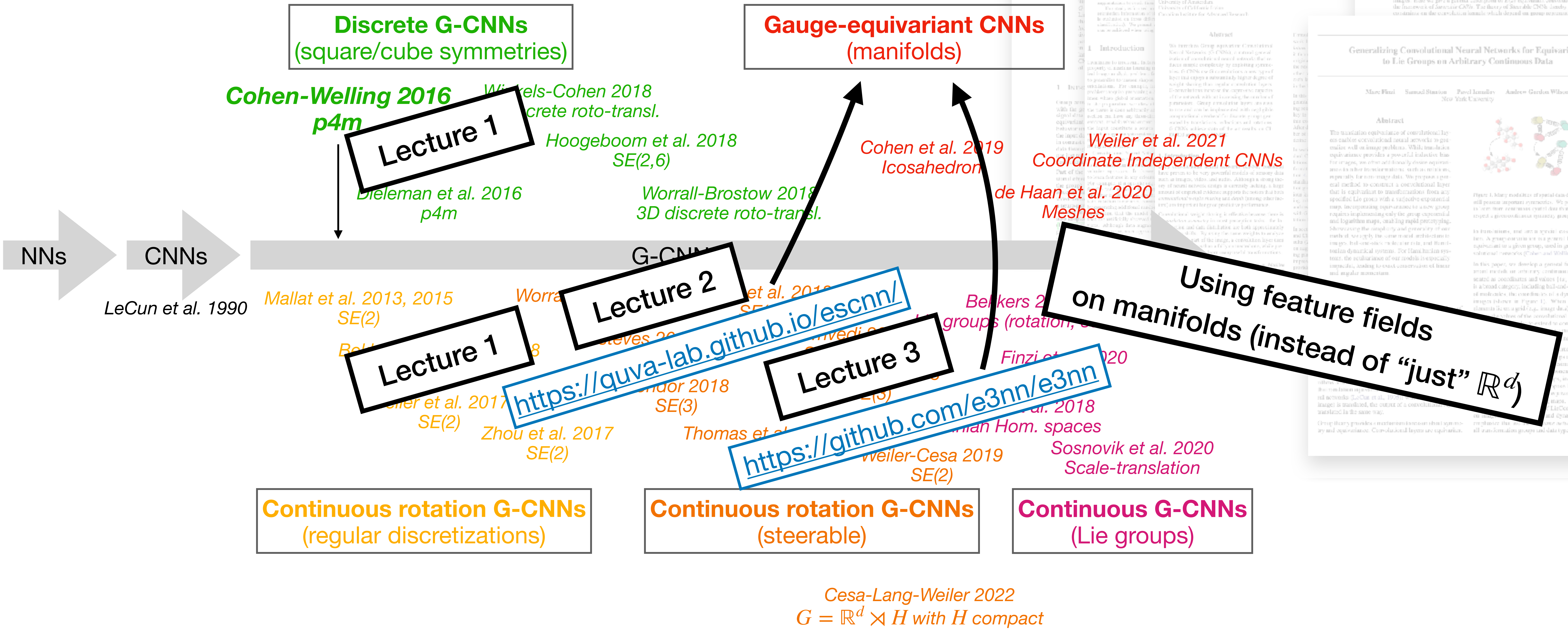
From plain NNs to Gauge-equivariant CNNs

https://github.com/Chen-Cai-OSU/awesome-equivariant-network



From plain NNs to Gauge-equivariant CNNs

https://github.com/Chen-Cai-OSU/awesome-equivariant-network



Recall lecture 2.4

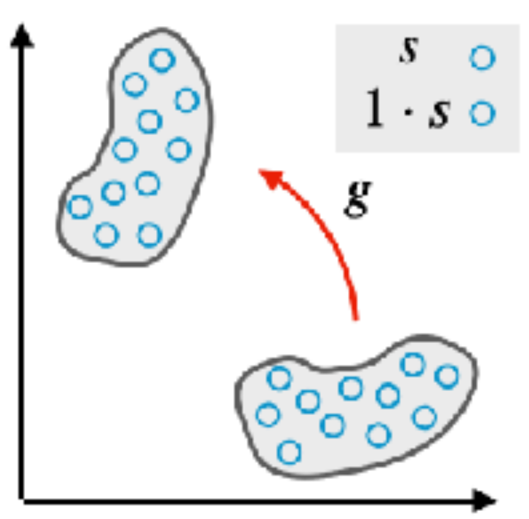
Feature field and induced representation

We call $\hat{f} : \mathbb{R}^d \rightarrow \mathbb{R}^{d_\rho}$ a feature vector field, or simply a **feature field**, if its

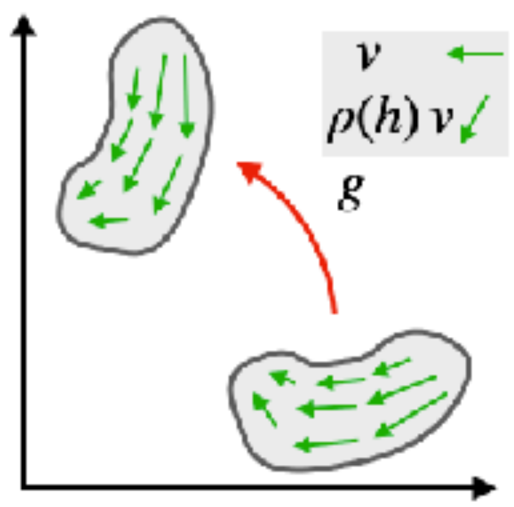
codomain	transforms via a *representation*	$\rho(h)$	of H
domain	transforms via the action	g^{-1}	of $G = (\mathbb{R}^d, +) \rtimes H$

Representation ρ defines the **type** of the field, and together with the group action of $G = (\mathbb{R}^d, +) \rtimes H$ defines the **induced representation**

$$\left(\mathrm{Ind}_H^G[\rho](\mathbf{x}, h)\hat{f}\right)(\mathbf{x}') := \rho(h)\hat{f}(h^{-1}(\mathbf{x}' - \mathbf{x}))$$



Scalar field ($\rho(h) = 1$)



vector field

Figure adapted from: Weiler, M., & Cesa, G. (2019). General e (2)-equivariant steerable cnns. NeurIPS
See also https://github.com/QUVA-Lab/e2cnn

2

Recall lecture 2.4

Feature field and induced representation

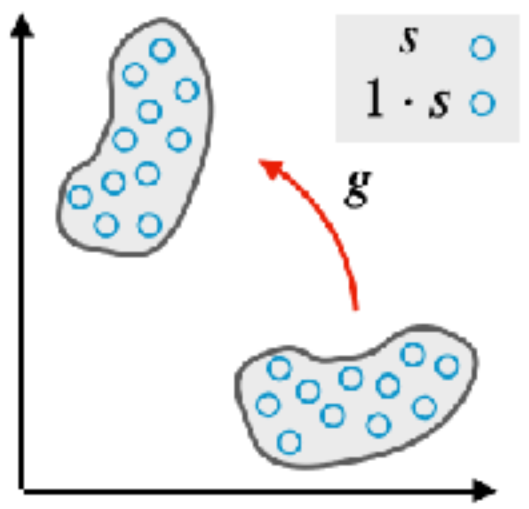
We call $\hat{f} : \mathbb{R}^d \rightarrow \mathbb{R}^{d_\rho}$ a feature vector field, c

codomain transforms via a represen

domain transforms via the action

Representation ρ defines the **type** of the field
induced representation

$$\left(\text{Ind}_H^G[\rho](\mathbf{x},\right.$$

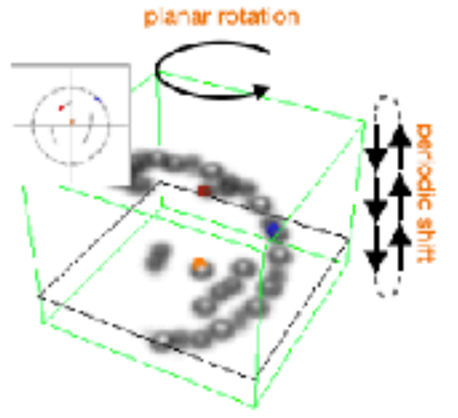


Scalar field ($\rho(h) = 1$)

Feature field and induced representation

Regular G feature maps: $f(\mathbf{x}, h)$ considered so far can be considered feature fields.

$$(\mathscr{L}_g f)(\mathbf{x}', h') = f(h^{-1}(\mathbf{x}' - \mathbf{x}), h^{-1}h)$$

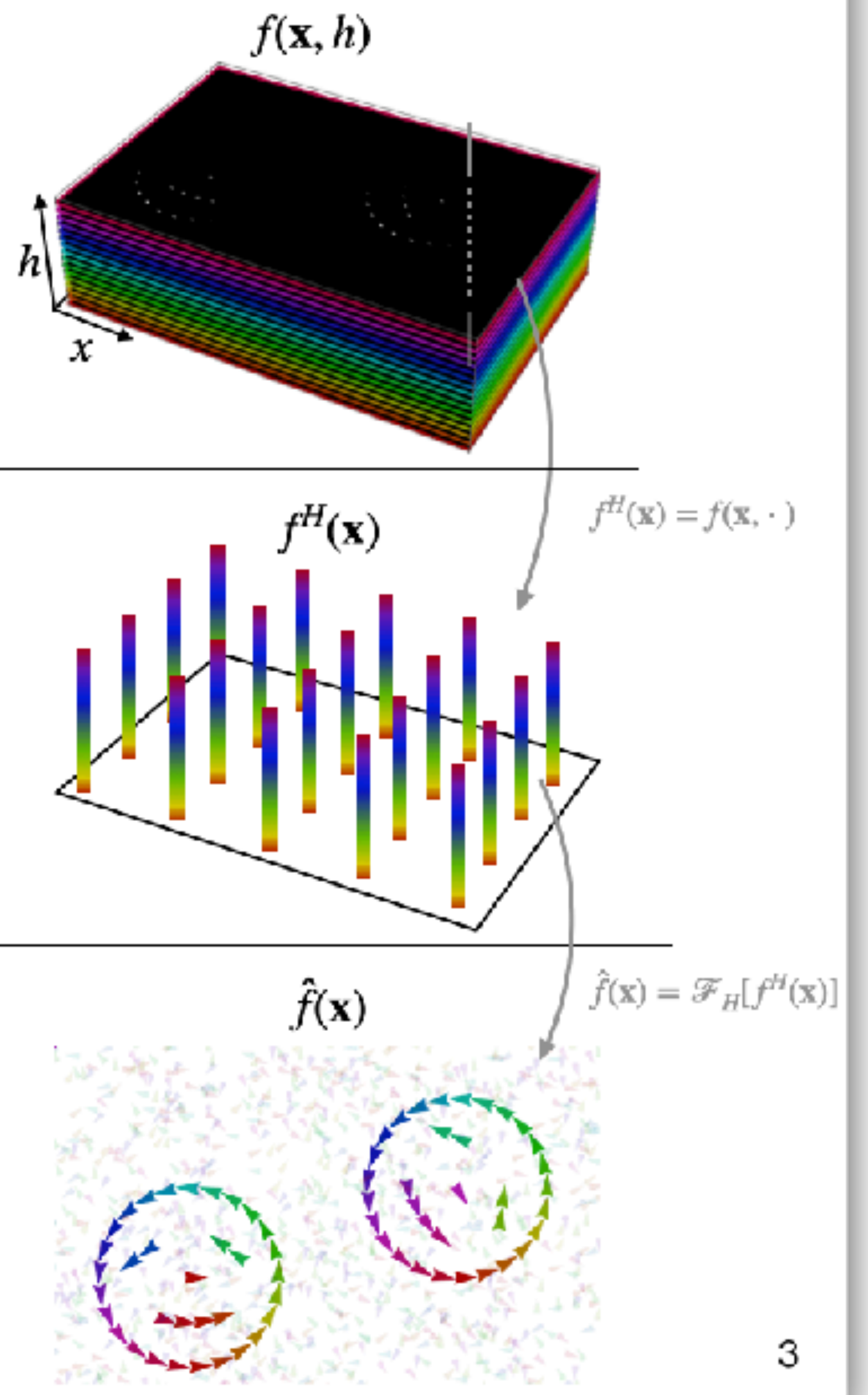


Regular H feature fields: Let $f^H(\mathbf{x}) = f(\mathbf{x}, \cdot\,)$ be the field of functions $f^H(\mathbf{x}) : H \rightarrow \mathbb{R}$ on the subgroup H, then the functions (**fibers**) transform via the regular representation $\mathscr{L}_h^H$ (recall. $\mathscr{L}_h^H f(h') = f(h^{-1}h')$)

$$(\mathscr{L}_g f)(\mathbf{x}', h') \iff (\text{Ind}_H^G[\ \mathscr{L}_h^H\ ](\mathbf{x}, h) f^H)(\mathbf{x}')$$

Steerable H feature fields: Since the fibers $f^H(\mathbf{x})$ are functions on H we can represent them via their Fourier coefficients $\hat{f}(\mathbf{x}) = \mathscr{F}_H[f^H(\mathbf{x})]$. These vectors of coefficients transform via irreps $\rho(h) = \bigoplus_l \rho_l(h)$

$$(\mathscr{L}_g f)(\mathbf{x}', h') \iff \left(\text{Ind}_H^G[\ \mathscr{L}_h^H\ ](\mathbf{x}, h)\hat{f}\right)(\mathbf{x}') \iff \left(\text{Ind}_H^G[\ \rho(h)\ ](\mathbf{x}, h)\hat{f}\right)(\mathbf{x}')$$



Recall lecture 2.4

Feature field and induced representation

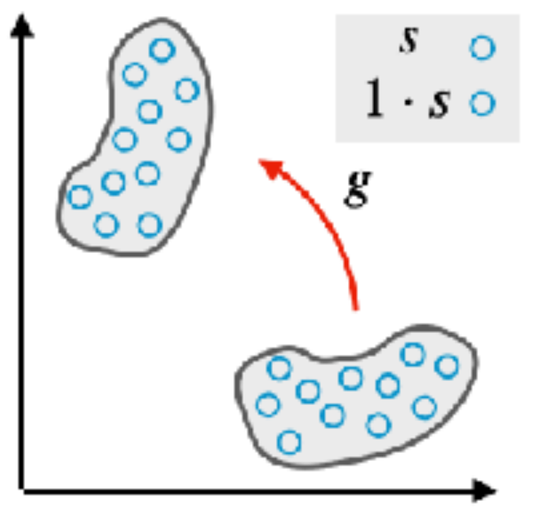
We call $\hat{f}: \mathbb{R}^d \rightarrow \mathbb{R}^{d_\rho}$ a feature vector field, c

codomain transforms via a represen

domain transforms via the action

Representation ρ defines the **type** of the field
induced representation

$$\left(\text{Ind}_H^G[\rho](\mathbf{x},\right.$$

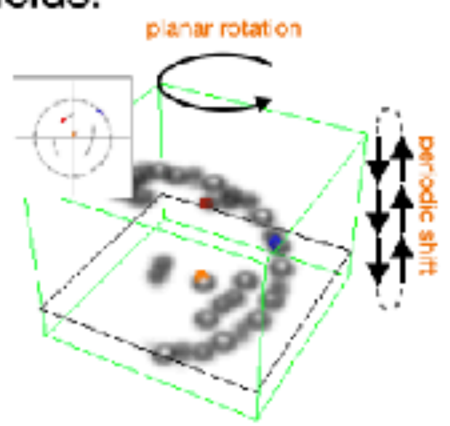


Scalar field ($\rho(h) = 1$)

Feature field and induced representation

Regular G feature maps: $f(\mathbf{x}, h)$ considered so far can be considered feature fields.

$$(\mathscr{L}_g f)(\mathbf{x}', h') = f(h^{-1}(\mathbf{x}' - \mathbf{x}), h^{-1}h)$$



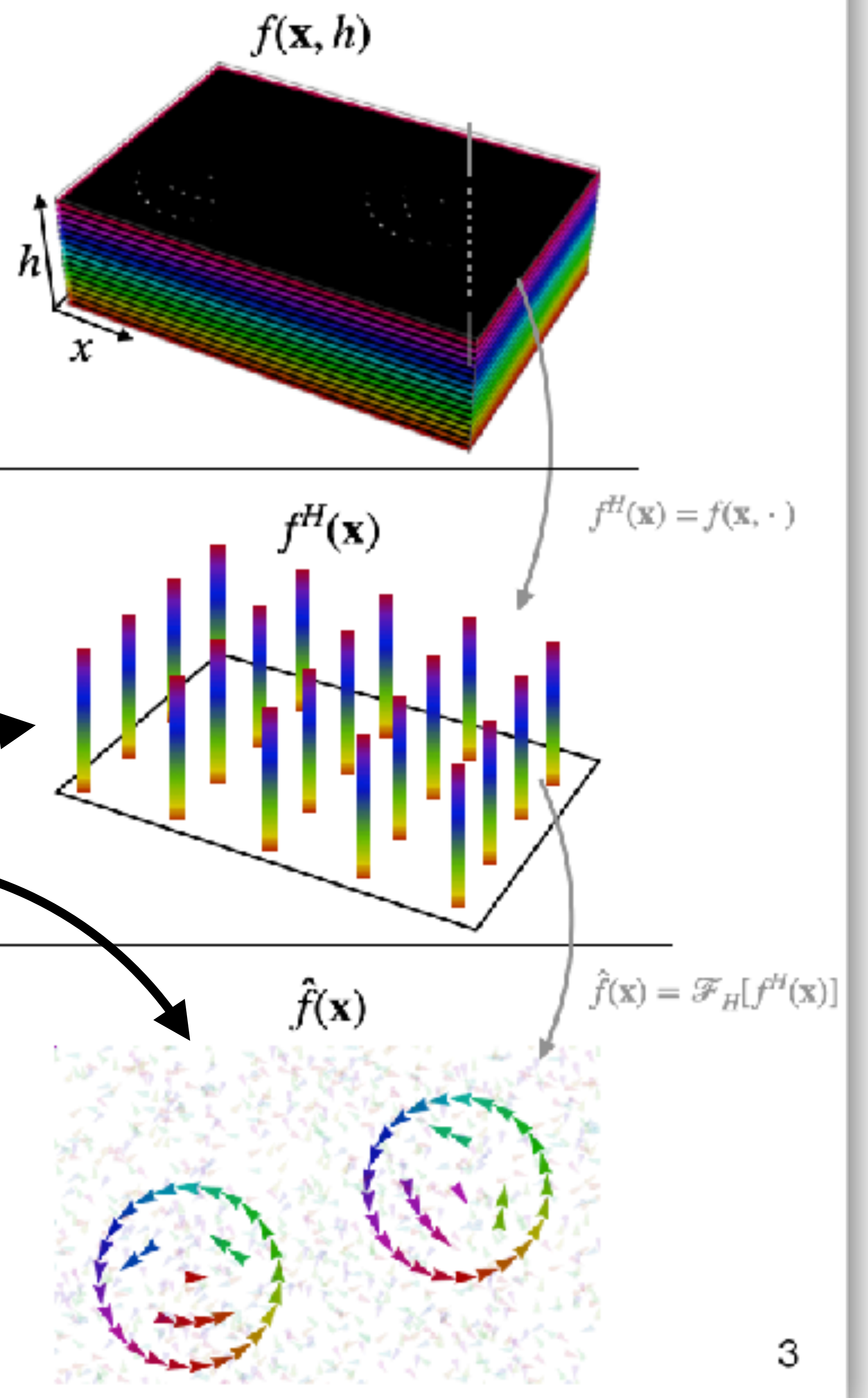
Regular H feature fields: Let $f^H(\mathbf{x}) = f(\mathbf{x}, \cdot\,)$ be the field of functions $f^H(\mathbf{x}) : H \rightarrow \mathbb{R}$ on the subgroup H, then the functions (**fibers**) transform via the regular representation $\mathscr{L}_h^H$ (recall. $\mathscr{L}_h^H f(h') = f(h^{-1}h')$)

$$(\mathscr{L}_g f)(\mathbf{x}', h' \ldots \mathscr{L}_h^H](\mathbf{x}, h) f^H)(\mathbf{x}')$$

Fibers (with oriented features) given relative to globally shared reference frame

Steerable H feature fields: Since the fibers $f^H(\mathbf{x})$ are functions … er coefficients $\hat{f}(\mathbf{x}) = \mathscr{F}_H[f^H(\mathbf{x})]$. These vectors of coefficients transform via irreps $\rho(h)$

$$(\mathscr{L}_g f)(\mathbf{x}', h') \iff \left(\text{Ind}_H^G[\ \mathscr{L}_h^H\ ](\mathbf{x}, h)\hat{f}\right)(\mathbf{x}') \iff \left(\text{Ind}_H^G[\ \rho(h)\ ](\mathbf{x}, h)\hat{f}\right)(\mathbf{x}')$$



Coordinate Independent Convolutional Networks

Isometry and Gauge Equivariant Convolutions on Riemannian Manifolds

Maurice Weiler
University of Amsterdam
m.weiler.ml@gmail.com

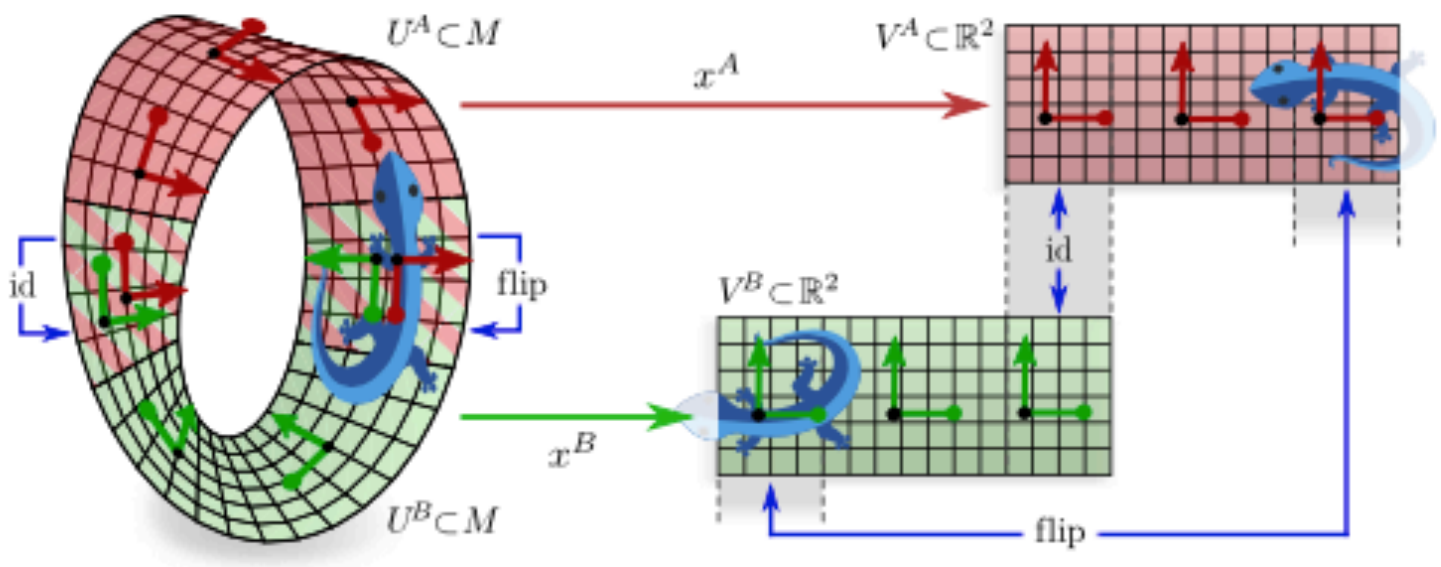
Patrick Forré
University of Amsterdam
p.d.forre@uva.nl

Erik Verlinde
University of Amsterdam
e.p.verlinde@uva.nl

Max Welling
University of Amsterdam,
Qualcomm AI Research
m.welling@uva.nl

Abstract

Motivated by the vast success of deep convolutional networks, there is a great interest in generalizing convolutions to non-Euclidean manifolds. A major complication in comparison to flat spaces is that it is unclear in which alignment a convolution kernel should be applied on a manifold. The underlying reason for this ambiguity is that general manifolds do not come with a canonical choice of reference frames (gauge). Kernels and features therefore have to be expressed relative to *arbitrary coordinates*. We argue that the particular choice of coordinatization should not affect a network's inference – it should be *coordinate independent*. A simultaneous demand for coordinate independence and weight sharing is shown to result in a requirement on the network to be *equivariant under local gauge transformations* (changes of local reference frames). The ambiguity of reference frames depends thereby on the *G-structure* of the manifold, such that the necessary level of gauge equivariance is prescribed by the corresponding *structure group G*. Coordinate independent convolutions are proven to be equivariant w.r.t. those *isometries* that are symmetries of the G-structure. The resulting theory is formulated in a coordinate free fashion in terms of fiber bundles. To exemplify the design of coordinate independent convolutions, we implement a convolutional network on the Möbius strip. The generality of our differential geometric formulation of convolutional networks is demonstrated by an extensive literature review which explains a large number of Euclidean CNNs, spherical CNNs and CNNs on general surfaces as specific instances of coordinate independent convolutions.



Gauge Equivariant Convolutional Networks and the Icosahedral CNN

Taco S. Cohen [*1] **Maurice Weiler** [*2] **Berkay Kicanaoglu** [*2] **Max Welling** [1]

Abstract

The principle of *equivariance to symmetry transformations* enables a theoretically grounded approach to neural network architecture design. Equivariant networks have shown excellent performance and data efficiency on vision and medical imaging problems that exhibit symmetries. Here we show how this principle can be extended beyond global symmetries to local gauge transformations. This enables the development of a very general class of convolutional neural networks on manifolds that depend only on the intrinsic geometry, and which includes many popular methods from equivariant and geometric deep learning.

We implement gauge equivariant CNNs for signals defined on the surface of the icosahedron, which provides a reasonable approximation of the sphere. By choosing to work with this very regular manifold, we are able to implement the gauge equivariant convolution using a single conv2d call, making it a highly scalable and practical alternative to Spherical CNNs. Using this method, we demonstrate substantial improvements over previous methods on the task of segmenting omnidirectional images and global climate patterns.

1. Introduction

By and large, progress in deep learning has been achieved through intuition-guided experimentation. This approach is indispensable and has led to many successes, but has not produced a deep understanding of *why and when* certain architectures work well. As a result, every new application requires an extensive architecture search, which comes at a significant labor and energy cost.

[*]Equal contribution [1]Qualcomm AI Research, Amsterdam, NL. [2]Qualcomm-University of Amsterdam (QUVA) Lab. Theory co-developed by Cohen & Weiler. Correspondence to: Taco S. Cohen <taco.cohen@gmail.com>, Maurice Weiler <m.weiler@uva.nl>.

Proceedings of the 36th International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019.

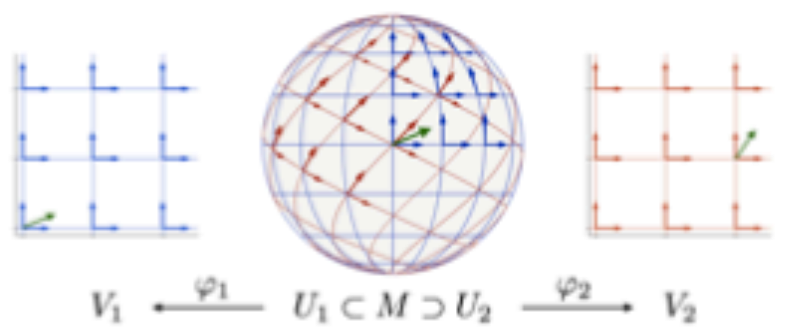


Figure 1. A gauge is a smoothly varying choice of tangent frame on a subset U of a manifold M. A gauge is needed to represent geometrical quantities such as convolutional filters and feature maps (i.e. fields), but the choice of gauge is ultimately arbitrary. Hence, the network should be equivariant to gauge transformations, such as the change between red and blue gauge pictured here.

Although a theory that tells us which architecture to use for any given problem is clearly out of reach, we can nevertheless come up with *general principles* to guide architecture search. One such rational design principle that has met with substantial empirical success (Winkels & Cohen, 2018; Zaheer et al., 2017; Lunter & Brown, 2018) maintains that network architectures should be equivariant to symmetries.

Besides the ubiquitous translation equivariant CNN, equivariant networks have been developed for sets, graphs, and homogeneous spaces like the sphere (see Sec. 3). In each case, the network is made equivariant to the global symmetries of the underlying space. However, manifolds do not in general have global symmetries, and so it is not obvious how one might develop equivariant CNNs for them.

General manifolds do however have *local gauge symmetries*, and as we will show in this paper, taking these into account is not just useful but *necessary* if one wishes to build manifold CNNs that depend only on the intrinsic geometry. To this end, we define a convolution-like operation on general manifolds M that is equivariant to local gauge transformations (Fig. 1). This *gauge equivariant convolution* takes as input a number of *feature fields* on M of various types (analogous to matter fields in physics), and produces as output new feature fields. Each field is represented by a number of feature maps, whose activations are interpreted as the coefficients of a geometrical object (e.g. scalar, vector, tensor, etc.) relative to a spatially varying frame (i.e. gauge). The network is constructed such that if the gauge is changed,

GAUGE EQUIVARIANT MESH CNNS
ANISOTROPIC CONVOLUTIONS ON GEOMETRIC GRAPHS

Pim de Haan[*]
Qualcomm AI Research[†]
University of Amsterdam

Maurice Weiler[*]
QUVA Lab
University of Amsterdam

Taco Cohen
Qualcomm AI Research

Max Welling
Qualcomm AI Research
University of Amsterdam

ABSTRACT

A common approach to define convolutions on meshes is to interpret them as a graph and apply graph convolutional networks (GCNs). Such GCNs utilize *isotropic* kernels and are therefore insensitive to the relative orientation of vertices and thus to the geometry of the mesh as a whole. We propose Gauge Equivariant Mesh CNNs which generalize GCNs to apply *anisotropic* gauge equivariant kernels. Since the resulting features carry orientation information, we introduce a geometric message passing scheme defined by parallel transporting features over mesh edges. Our experiments validate the significantly improved expressivity of the proposed model over conventional GCNs and other methods.

1 INTRODUCTION

Convolutional neural networks (CNNs) have been established as the default method for many machine learning tasks like speech recognition or planar and volumetric image classification and segmentation. Most CNNs are restricted to flat or spherical geometries, where convolutions are easily defined and optimized implementations are available. The empirical success of CNNs on such spaces has generated interest to generalize convolutions to more general spaces like graphs or Riemannian manifolds, creating a field now known as geometric deep learning (Bronstein et al., 2017).

A case of specific interest is convolution on *meshes*, the discrete analog of 2-dimensional embedded Riemannian manifolds. Mesh CNNs can be applied to tasks such as detecting shapes, registering different poses of the same shape and shape segmentation. If we forget the positions of vertices, and which vertices form faces, a mesh M can be represented by a graph $\mathcal{G}$. This allows for the application of *graph convolutional networks* (GCNs) to processing signals on meshes.

[*]Equal Contribution
[†]Qualcomm AI Research is an initiative of Qualcomm Technologies, Inc.

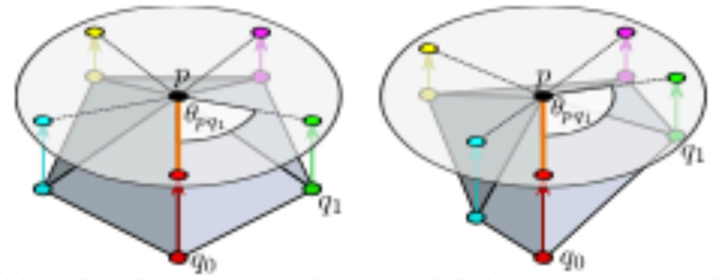


Figure 1: Two local neighbourhoods around vertices p and their representations in the tangent planes T_pM. The distinct geometry of the neighbourhoods is reflected in the different angles θ_{pq_i} of incident edges from neighbours q_i. Graph convolutional networks apply isotropic kernels and can therefore not distinguish both neighbourhoods. Gauge Equivariant Mesh CNNs apply anisotropic kernels and are therefore sensitive to orientations. The arbitrariness of reference orientations, determined by a choice of neighbour q_0, is accounted for by the gauge equivariance of the model.

Coordinate Independent Convolutional Networks

Isometry and Gauge Equivariant Convolutions on Riemannian Manifolds

Maurice Weiler
University of Amsterdam
m.weiler.ml@gmail.com

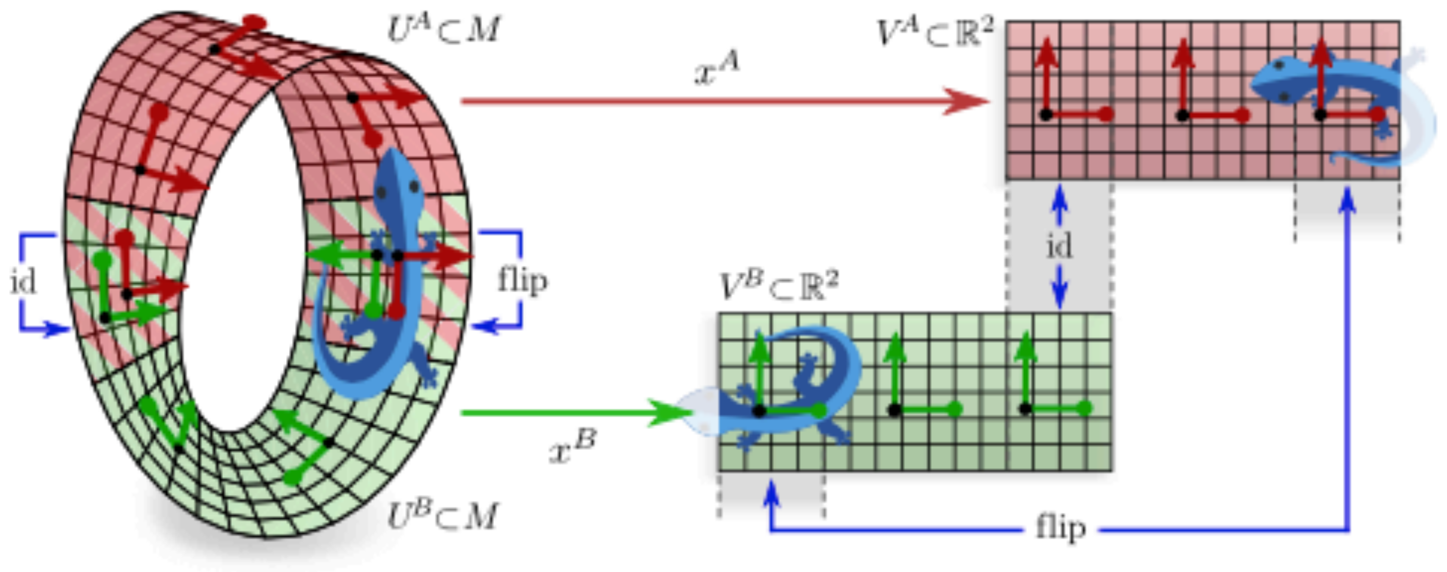
Patrick Forré
University of Amsterdam
p.d.forre@uva.nl

Erik Verlinde
University of Amsterdam
e.p.verlinde@uva.nl

Max Welling
University of Amsterdam,
Qualcomm AI Research
m.welling@uva.nl

Abstract

Motivated by the vast success of deep convolutional networks, there is a great interest in generalizing convolutions to non-Euclidean manifolds. A major complication in comparison to flat spaces is that it is unclear in which alignment a convolution kernel should be applied on a manifold. The underlying reason for this ambiguity is that general manifolds do not come with a canonical choice of reference frames (gauge). Kernels and features therefore have to be expressed relative to *arbitrary coordinates*. We argue that the particular choice of coordinatization should not affect a network's inference – it should be *coordinate independent*. A simultaneous demand for coordinate independence and weight sharing is shown to result in a requirement on the network to be *equivariant under local gauge transformations* (changes of local reference frames). The ambiguity of reference frames depends thereby on the *G-structure* of the manifold, such that the necessary level of gauge equivariance is prescribed by the corresponding *structure group* G. Coordinate independent convolutions are proven to be equivariant w.r.t. those *isometries* that are symmetries of the G-structure. The resulting theory is formulated in a coordinate free fashion in terms of fiber bundles. To exemplify the design of coordinate independent convolutions, we implement a convolutional network on the Möbius strip. The generality of our differential geometric formulation of convolutional networks is demonstrated by an extensive literature review which explains a large number of Euclidean CNNs, spherical CNNs and CNNs on general surfaces as specific instances of coordinate independent convolutions.



Coordinate Independent Convolutional Networks

Isometry and Gauge Equivariant Convolutions

Maurice Weiler
University of Amsterdam
m.weiler.ml@gmail.com

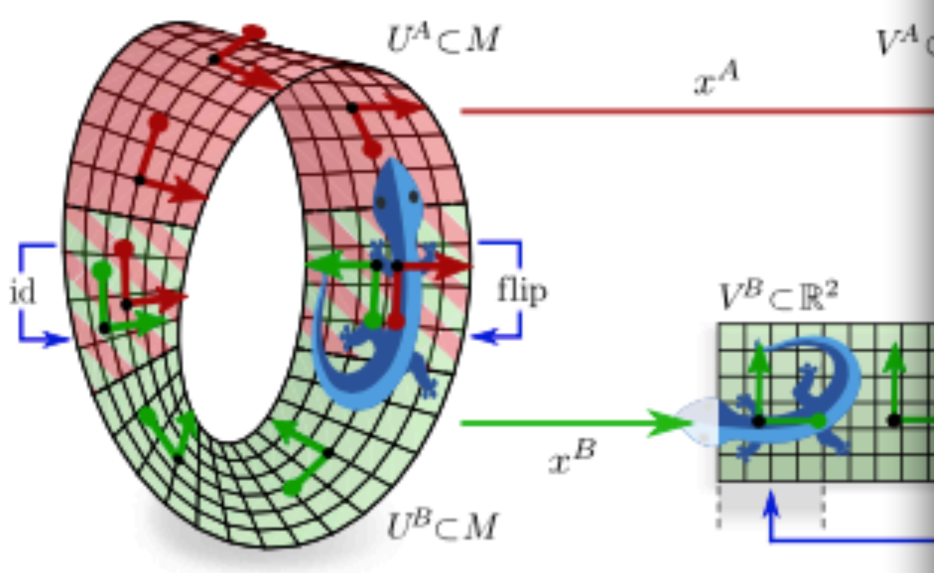
Pa
Universi
p.d.

Erik Verlinde
University of Amsterdam
e.p.verlinde@uva.nl

Ma
Universit
Qualcom
m.wel

Abstract

Motivated by the vast success of deep convolutional netwo
generalizing convolutions to non-Euclidean manifolds. A m
ison to flat spaces is that it is unclear in which alignment a
applied on a manifold. The underlying reason for this ambig
do not come with a canonical choice of reference frames (
therefore have to be expressed relative to *arbitrary coordi*
ticular choice of coordinatization should not affect a netw
coordinate independent. A simultaneous demand for coordi
sharing is shown to result in a requirement on the network
gauge transformations (changes of local reference frames)
frames depends thereby on the *G-structure* of the manifold
of gauge equivariance is prescribed by the corresponding *s*
independent convolutions are proven to be equivariant w.r.t.
metries of the G-structure. The resulting theory is formulate
in terms of fiber bundles. To exemplify the design of coordin
we implement a convolutional network on the Möbius strip
ential geometric formulation of convolutional networks is
literature review which explains a large number of Euclide
CNNs on general surfaces as specific instances of coordinat



	manifold M	structure group G	global symmetry Aff_{GM} or Isom_{GM}	representation ρ	citation
1	$\mathbb{E}_d$	$\{e\}$	$\mathcal{T}_d$	trivial	[130, 253] + any conventional CNN
2	$\mathbb{E}_1$	$\mathcal{S}$	$\mathcal{T}_1 \rtimes \mathcal{S}$	regular	[186]
3		$\mathcal{R}$	$\mathcal{T}_2 \rtimes \mathcal{R}$	regular	[234]
4				irreps	[244, 234, 231]
5		SO(2)	SE(2)	regular	[51, 33, 257, 34, 236, 8, 93, 192] [234, 79, 125, 210, 232, 185, 158] [201, 7, 67, 227, 83, 159, 231, 92] [50, 206, 19, 207, 208, 164, 29, 86]
6				quotients	[34, 234]
7				regular $\xrightarrow{\text{pool}}$ trivial	[33, 143, 234]
8	$\mathbb{E}_2$			regular $\xrightarrow{\text{pool}}$ vector	[144, 234]
9				trivial	[110, 234]
10				irreps	[234]
11		O(2)	E(2)	regular	[51, 33, 93, 34, 234] [159, 79, 201]
12				quotients	[34]
13				regular $\xrightarrow{\text{pool}}$ trivial	[234]
14				induced SO(2)-irreps	[234]
15		$\mathcal{S}$	$\mathcal{T}_2 \rtimes \mathcal{S}$	regular	[243, 212, 7, 258]
16				regular $\xrightarrow{\text{pool}}$ trivial	[77]
17				irreps	[235, 224, 156, 120, 2, 6]
18		SO(3)	SE(3)	quaternion	[250]
19				regular	[67, 241, 242]
20				regular $\xrightarrow{\text{pool}}$ trivial	[3]
21				regular	[241]
22	$\mathbb{E}_3$	O(3)	E(3)	quotient O(3)/O(2)	[103]
23				irrep $\xrightarrow{\text{norm}}$ trivial	[174]
24		C_4	$\mathcal{T}_3 \rtimes C_4$	regular	[219]
25		D_4	$\mathcal{T}_3 \rtimes D_4$	regular	[219]
26	$\mathbb{E}_{d-1,1}$	SO(d−1, 1)	$\mathcal{T}_d \rtimes$ SO(d−1, 1)	irreps	[205]
27	$\mathbb{E}_2 \setminus \{0\}$	$\{e\}$	SO(2)	trivial	[30, 67]
28			SO(2) $\times \mathcal{S}$		[62, 67]
29	$\mathbb{E}_3 \setminus \{0\}$	O(2)	O(3)	trivial	[178]
30		$\{e\}$	$\{e\}$	trivial	[13]
31		SO(2)	SO(3)	irreps	[122, 64]
32	S^2			regular	[35, 111]
33		O(2)	O(3)	trivial	[61, 169, 245]
34	$S^2 \setminus$ poles	$\{e\}$	SO(2)	trivial	[39, 222, 254, 149, 105] [217, 218, 55, 131]
35	icosahedron	C_6	I ($\approx$ SO(3))	regular	[38]
36	ico \ poles	$\{e\}$	C_5 ($\approx$ SO(2))	trivial	[251, 139]
37				irreps	[238]
38	surface (d=2) (e.g. meshes)	SO(2)	$\mathrm{Isom}_+(M)$	regular	[173, 220, 246, 48]
39				regular $\xrightarrow{\text{pool}}$ trivial	[150, 151, 160, 220]
40		D_4	$\mathrm{Isom}_{D_4 M}$	trivial	[98]
41		$\{e\}$	$\mathrm{Isom}_{\{e\}M}$	trivial	[160, 194, 106, 221, 133]
42	Möbius strip	$\mathcal{R}$	SO(2)	irreps	Section 5
43				regular	

Table 6: Classification of convolutional networks in the literature from the viewpoint of coordinate independent CNNs. Bold lines separate different geometries. The affine group equivariant convolutions on Euclidean spaces $\mathbb{E}_d$ (rows 1-26) are reviewed in Section 9. Section 10 discusses GM-convolutions on punctured Euclidean spaces $\mathbb{E}_d \setminus \{0\} \cong S^{d-1} \times \mathbb{R}^+$ (rows 27-30). Details on spherical CNNs (rows 31-36) are found in Section 11. The models in rows (37-41) operate on general surfaces, mostly represented by triangle meshes; see Section 12. The last two lines list our Möbius convolutions from Section 5. $\mathcal{T}_d$, $\mathcal{R}$ and $\mathcal{S}$ denote the translation, reflection and scaling group, respectively, while C_N and D_N are cyclic and dihedral groups. Infinite-dimensional representations are in implementations discretized or sampled. For instance, the regular representations of SO(2) or O(2) are typically approximated by the regular representations of cyclic or dihedral groups C_N or D_N.

Coordinate Independent Convolutional Networks

Isometry and Gauge Equivariant Convolutions

Maurice Weiler
University of Amsterdam
m.weiler.ml@gmail.com

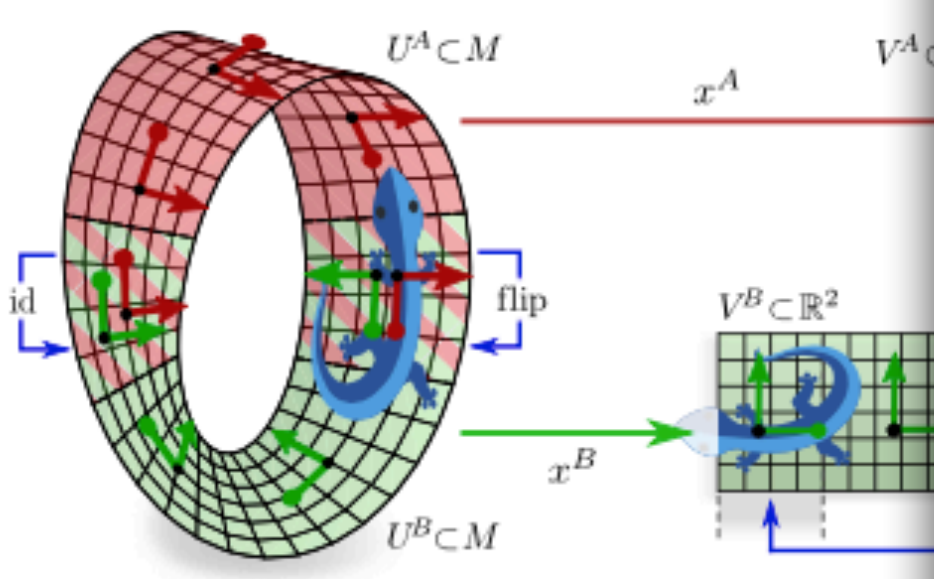
Pa
Univers
p.d.

Erik Verlinde
University of Amsterdam
e.p.verlinde@uva.nl

Ma
Universit
Qualcon
m.wel

Abstract

Motivated by the vast success of deep convolutional netwo
generalizing convolutions to non-Euclidean manifolds. A m
ison to flat spaces is that it is unclear in which alignment a
applied on a manifold. The underlying reason for this ambig
do not come with a canonical choice of reference frames (
therefore have to be expressed relative to *arbitrary coordi*
ticular choice of coordinatization should not affect a netw
coordinate independent. A simultaneous demand for coordi
sharing is shown to result in a requirement on the network
gauge transformations (changes of local reference frames)
frames depends thereby on the *G*-structure of the manifold
of gauge equivariance is prescribed by the corresponding *s*
independent convolutions are proven to be equivariant w.r.t.
metries of the *G*-structure. The resulting theory is formulate
in terms of fiber bundles. To exemplify the design of coordin
we implement a convolutional network on the Möbius strip
ential geometric formulation of convolutional networks is
literature review which explains a large number of Euclidea
CNNs on general surfaces as specific instances of coordinate



	manifold M	structure group G	global symmetry Aff_{GM} or Isom_{GM}	representation ρ	citation
1	$\mathbb{E}_d$	$\{e\}$	$\mathcal{T}_d$	trivial	[130, 253] + any conventional CNN
2	$\mathbb{E}_1$	$\mathcal{S}$	$\mathcal{T}_1 \rtimes \mathcal{S}$	regular	[186]
3		$\mathcal{R}$	$\mathcal{T}_2 \rtimes \mathcal{R}$	regular	[234]
4				irreps	[244, 234, 231]
5		SO(2)	SE(2)	regular	[51, 33, 257, 34, 236, 8, 93, 192, 234, 79, 125, 210, 232, 185, 158, 201, 7, 67, 227, 83, 159, 231, 92, 50, 206, 19, 207, 208, 164, 29, 86]
6				quotients	[34, 234]
7				regular $\xrightarrow{\text{pool}}$ trivial	[33, 143, 234]
8	$\mathbb{E}_2$			regular $\xrightarrow{\text{pool}}$ vector	[144, 234]
9				trivial	[110, 234]
10				irreps	[234]
11		O(2)	E(2)	regular	[51, 33, 93, 34, 234, 159, 79, 201]
12				quotients	[34]
13				regular $\xrightarrow{\text{pool}}$ trivial	[234]
14				induced SO(2)-irreps	[234]
15		$\mathcal{S}$	$\mathcal{T}_2 \rtimes \mathcal{S}$	regular	[243, 212, 7, 258]
16				regular $\xrightarrow{\text{pool}}$ trivial	[77]
17				irreps	[235, 224, 156, 120, 2, 6]
18		SO(3)	SE(3)	quaternion	[250]
19				regular	[67, 241, 242]
20				regular $\xrightarrow{\text{pool}}$ trivial	[3]
21				regular	[241]
22	$\mathbb{E}_3$	O(3)	E(3)	quotient O(3)/ O(2)	[103]
23				irrep $\xrightarrow{\text{norm}}$ trivial	[174]
24		C_4	$\mathcal{T}_3 \rtimes C_4$	regular	[219]
25		D_4	$\mathcal{T}_3 \rtimes D_4$	regular	[219]
26	$\mathbb{E}_{d-1,1}$	SO(d−1, 1)	$\mathcal{T}_d \rtimes$ SO(d−1, 1)	irreps	[205]
27	$\mathbb{E}_2 \backslash \{0\}$	$\{e\}$	SO(2)	trivial	[30, 67]
28			SO(2) × $\mathcal{S}$		[62, 67]
29	$\mathbb{E}_3 \backslash \{0\}$	O(2)	O(3)	trivial	[178]
30		$\{e\}$	$\{e\}$	trivial	[13]
31		SO(2)	SO(3)	irreps	[122, 64]
32	S^2			regular	[35, 111]
33		O(2)	O(3)	trivial	[61, 169, 245]
34	$S^2 \setminus$ poles	$\{e\}$	SO(2)	trivial	[39, 222, 254, 149, 105, 217, 218, 55, 131]
35	icosahedron	C_6	I (≈ SO(3))	regular	[38]
36	ico \ poles	$\{e\}$	C_5 (≈ SO(2))	trivial	[251, 139]
37				irreps	[238]
38	surface (d = 2) (e.g. meshes)	SO(2)	$\mathrm{Isom}_+(M)$	regular	[173, 220, 246, 48]
39				regular $\xrightarrow{\text{pool}}$ trivial	[150, 151, 160, 220]
40		D_4	$\mathrm{Isom}_{D_4 M}$	trivial	[98]
41		$\{e\}$	$\mathrm{Isom}_{\{e\}M}$	trivial	[160, 194, 106, 221, 133]
42	Möbius strip	$\mathcal{R}$	SO(2)	irreps	Section 5
43				regular	

Table 6: Classification of convolutional networks in the literature from the viewpoint of coordinate independent CNNs. Bold lines separate different geometries. The affine group equivariant convolutions on Euclidean spaces $\mathbb{E}_d$ (rows 1-26) are reviewed in Section 9. Section 10 discusses GM-convolutions on punctured Euclidean spaces $\mathbb{E}_d \backslash \{0\} \cong S^{d-1} \times \mathbb{R}^+$ (rows 27-30). Details on spherical CNNs (rows 31-36) are found in Section 11. The models in rows (37-41) operate on general surfaces, mostly represented by triangle meshes; see Section 12. The last two lines list our Möbius convolutions from Section 5. $\mathcal{T}_d$, $\mathcal{R}$ and $\mathcal{S}$ denote the translation, reflection and scaling group, respectively, while C_N and D_N are cyclic and dihedral groups. Infinite-dimensional representations are in implementations discretized or sampled. For instance, the regular representations of SO(2) or O(2) are typically approximated by the regular representations of cyclic or dihedral groups C_N or D_N.

Geodesic Convolutional Neural Networks on Riemannian Manifolds

Jonathan Masci[†*] Davide Boscaini[†*] Michael M. Bronstein[†] Pierre Vandergheynst[‡]
[†]USI, Lugano, Switzerland [‡]EPFL, Lausanne, Switzerland

Abstract

Feature descriptors play a crucial role in a wide range of geometry analysis and processing applications, including shape correspondence, retrieval, and segmentation. In this paper, we introduce Geodesic Convolutional Neural Networks (GCNN), a generalization of the convolutional networks (CNN) paradigm to non-Euclidean manifolds. Our construction is based on a local geodesic system of polar coordinates to extract "patches", which are then passed through a cascade of filters and linear and non-linear operators. The coefficients of the filters and linear combination weights are optimization variables that are learned to minimize a task-specific cost function. We use GCNN to learn invariant shape features, allowing to achieve state-of-the-art performance in problems such as shape description, retrieval, and correspondence.

1. Introduction

Feature descriptors are ubiquitous tools in shape analysis. Broadly speaking, a *local* feature descriptor assigns to each point on the shape a vector in some multi-dimensional descriptor space representing the local structure of the shape around that point. A *global* descriptor describes the whole shape. Local feature descriptors are used in higher-level tasks such as establishing correspondence between shapes [35], shape retrieval [8], or segmentation [43]. Global descriptors are often produced by aggregating local descriptors e.g. using the bag-of-features paradigm. Descriptor construction is largely application dependent, and one typically tries to make the descriptor discriminative (capture the structures that are important for a particular application, e.g. telling apart two classes of shapes), robust (invariant to some class of transformations or noise), compact (low dimensional), and computationally-efficient.

Previous work Early works on shape descriptors such as spin images [19], shape distributions [34], and integral volume descriptors [32] were based on *extrinsic* structures that are invariant under Euclidean transformations. The following generation of shape descriptors used *intrinsic* structures such as geodesic distances [15] that are preserved by isometric deformations. The success of image descriptors such as SIFT [31], HOG [13], MSER [33], and shape context [2] has led to several generalizations thereof to non-Euclidean domains (see e.g. [49, 14, 24], respectively). The works [11, 28] on diffusion and spectral geometry have led to the emergence of intrinsic spectral shape descriptors that are *dense* and isometry-invariant by construction. Notable examples in this family include heat kernel signatures (HKS) [45] and wave kernel signatures (WKS) [1].

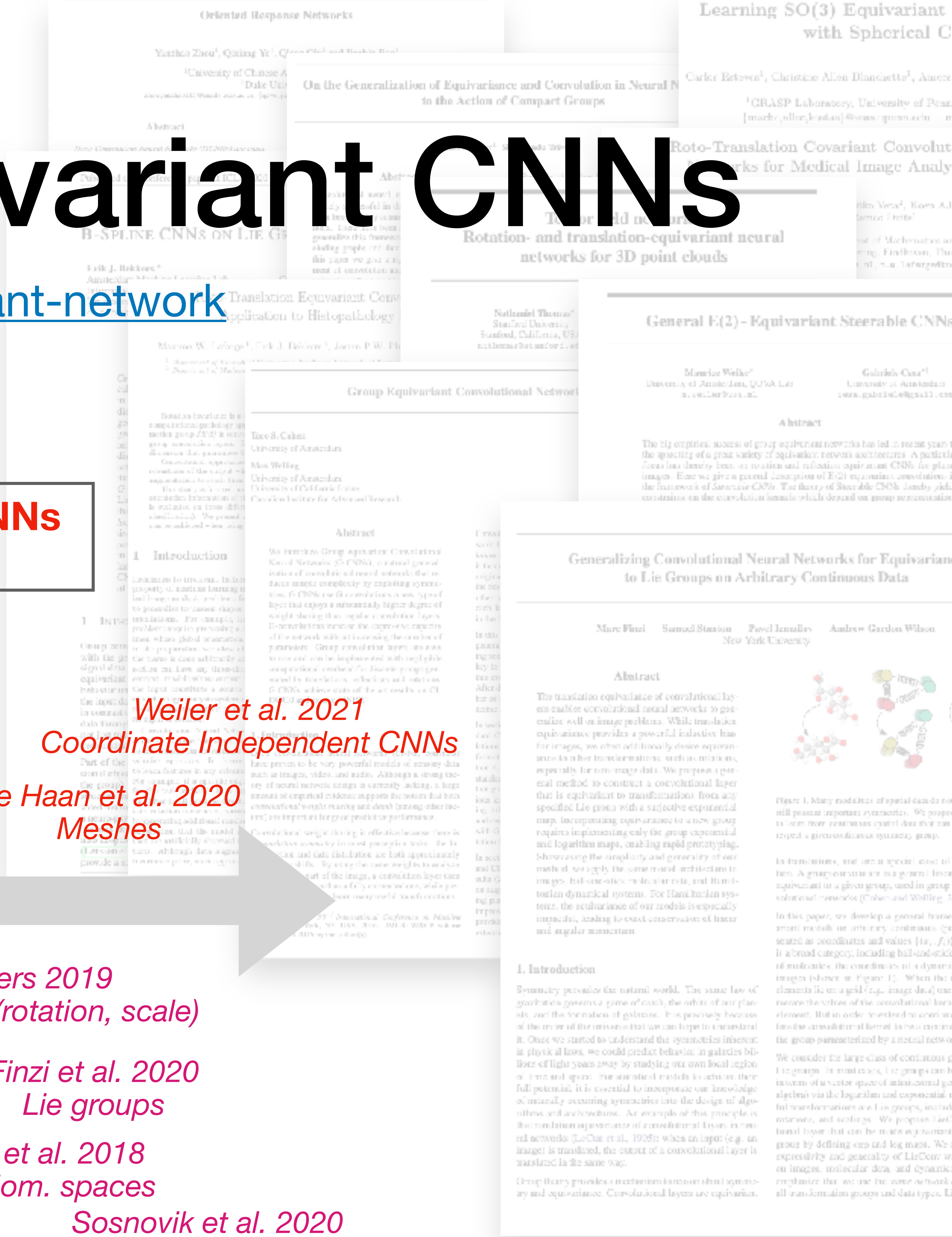
Arguing that in many cases it is hard to model invariance but rather easy to create examples of similar and dissimilar shapes, Litman and Bronstein [29] showed that HKS and WKS can be considered as particular parametric families of transfer functions applied to the Laplace-Beltrami operator eigenvalues and proposed to learn an optimal transfer function. Their work follows the recent trends in the image analysis domain, where hand-crafted descriptors are abandoned in favor of learning approaches. The past decade in computer vision research has witnessed the re-emergence of "deep learning" and in particular, convolutional neural network (CNN) techniques [17, 27], allowing to learn task-specific features from examples. CNNs achieve a breakthrough in performance in a wide range of applications such as image classification [26], segmentation [10], detection and localization [38, 42] and annotation [16, 21].

Learning methods have only recently started penetrating into the 3D shape analysis community in problems such as shape correspondence [39, 37], similarity [20], description [29, 47, 12], and retrieval [30]. CNNs have been applied to 3D data in the very recent works [48, 44] using standard (Euclidean) CNN architectures applied to volumetric 2D views shape representations, making them unsuitable for deformable shapes. Intrinsic versions of CNNs that would allow dealing with shape deformations are difficult to formulate due to the lack of shift invariance on Riemannian manifolds; we are aware of two recent works in that direction [9, 5].

Contribution In this paper, we propose Geodesic CNN (GCNN), an extension of the CNN paradigm to non-Euclidean manifolds based on local geodesic system of coor-

[*]equal contribution

From plain NNs to Gauge-equivariant CNNs



https://github.com/Chen-Cai-OSU/awesome-equivariant-network

Discrete G-CNNs
(square/cube symmetries)

Gauge-equivariant CNNs
(manifolds)

Cohen-Welling 2016
p4m

Winkels-Cohen 2018
3D discrete roto-transl.

Hoogeboom et al. 2018
SE(2,6)

Dieleman et al. 2016
p4m

Worrall-Brostow 2018
3D discrete roto-transl.

Cohen et al. 2019
Icosahedron

Weiler et al. 2021
Coordinate Independent CNNs

de Haan et al. 2020
Meshes

NNs → CNNs → G-CNNs

LeCun et al. 1990

Mallat et al. 2013, 2015
SE(2)

Bekkers et al. 2014-2018
SE(2)

Weiler et al. 2017
SE(2)

Zhou et al. 2017
SE(2)

Worrall et al. 2017
SE(2)

Esteves 2017
SO(3)

Kondor 2018
SE(3)

Thomas et al. 2018
SE(3)

Cohen et al. 2018
SE(3)

Kondor-Trivedi 2018
SE(3)

Weiler et al. 2018
SE(3)

Weiler-Cesa 2019
SE(2)

Bekkers 2019
Lie groups (rotation, scale)

Finzi et al. 2020
Lie groups

Chakraborty et al. 2018
Riemannian Hom. spaces

Sosnovik et al. 2020
Scale-translation

Continuous rotation G-CNNs
(regular discretizations)

Continuous rotation G-CNNs
(steerable)

Continuous G-CNNs
(Lie groups)

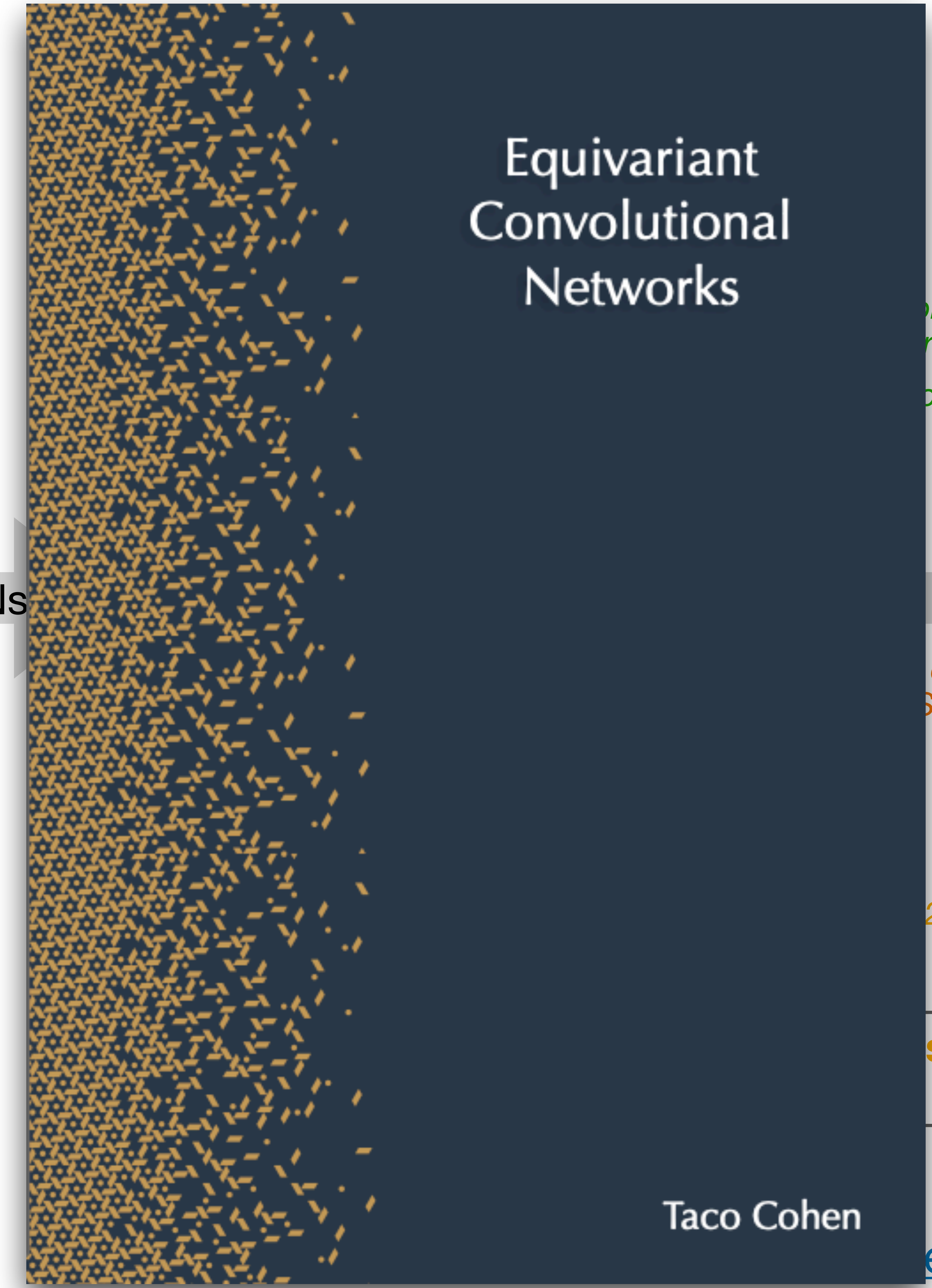
Cesa-Lang-Weiler 2022
$G = \mathbb{R}^d \rtimes H$ *with* H *compact*

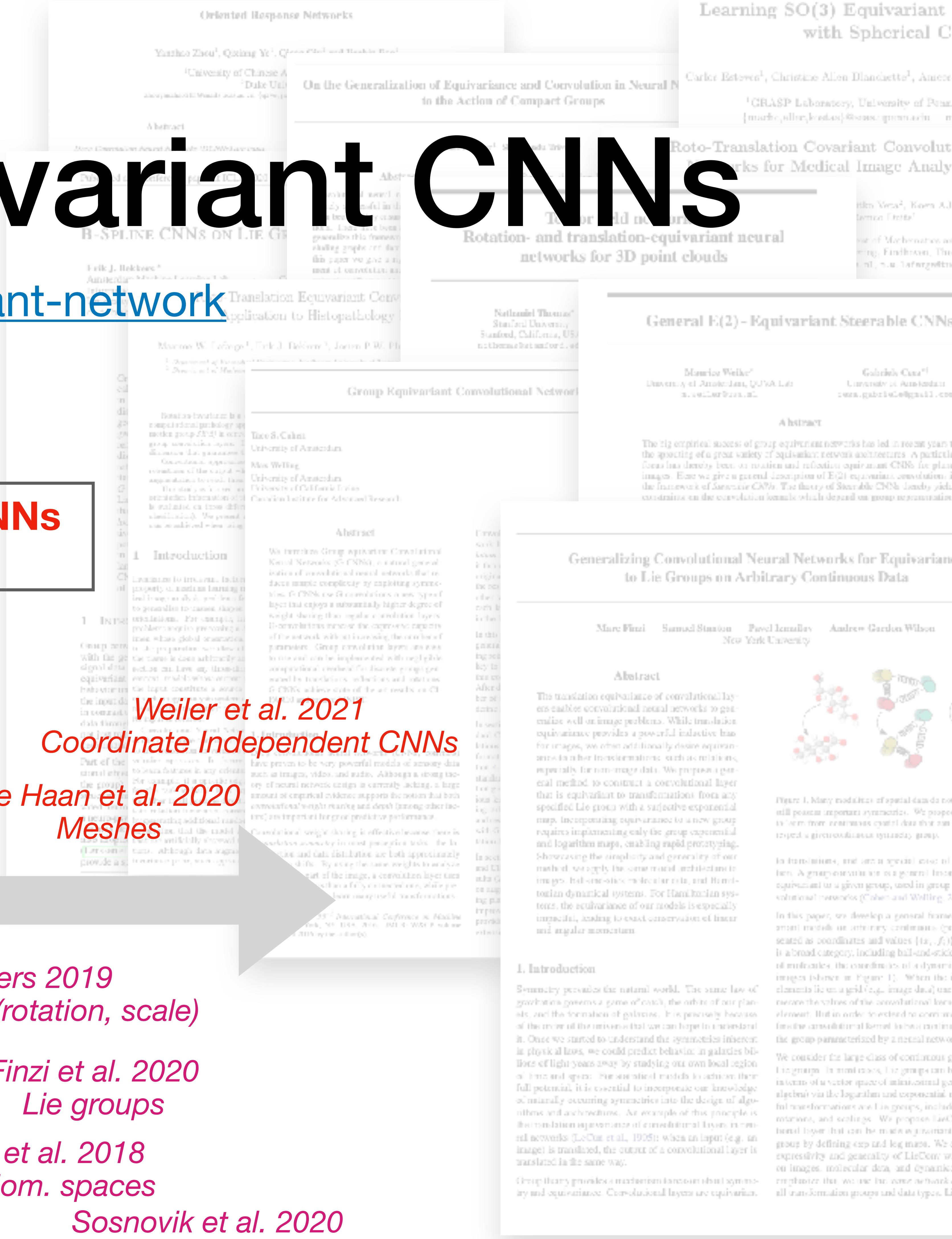
https://github.com/e3nn/e3nn

https://quva-lab.github.io/escnn/

From plain NNs to Gauge-equivariant CNNs

https://github.com/Chen-Cai-OSU/awesome-equivariant-network





NNs

G-CNNs

Gauge-equivariant CNNs
(manifolds)

...hen 2018
...oto-transl.

...ogeboom et al. 2018
SE(2,6)

Worrall-Brostow 2018
3D discrete roto-transl.

Cohen et al. 2019
Icosahedron

Weiler et al. 2021
Coordinate Independent CNNs

de Haan et al. 2020
Meshes

...et al. 2017
...E(2)

Esteves 2017
SO(3)

Kondor 2018
SE(3)

...2017

Cohen et al. 2018
SE(3)

Kondor-Trivedi 2018
SE(3)

Weiler et al. 2018
SE(3)

Thomas et al. 2018
SE(3)

Weiler-Cesa 2019
SE(2)

Bekkers 2019
Lie groups (rotation, scale)

Finzi et al. 2020
Lie groups

Chakraborty et al. 2018
Riemannian Hom. spaces

Sosnovik et al. 2020
Scale-translation

...s

Continuous rotation G-CNNs
(steerable)

Continuous G-CNNs
(Lie groups)

Cesa-Lang-Weiler 2022
$G = \mathbb{R}^d \rtimes H$ with H compact

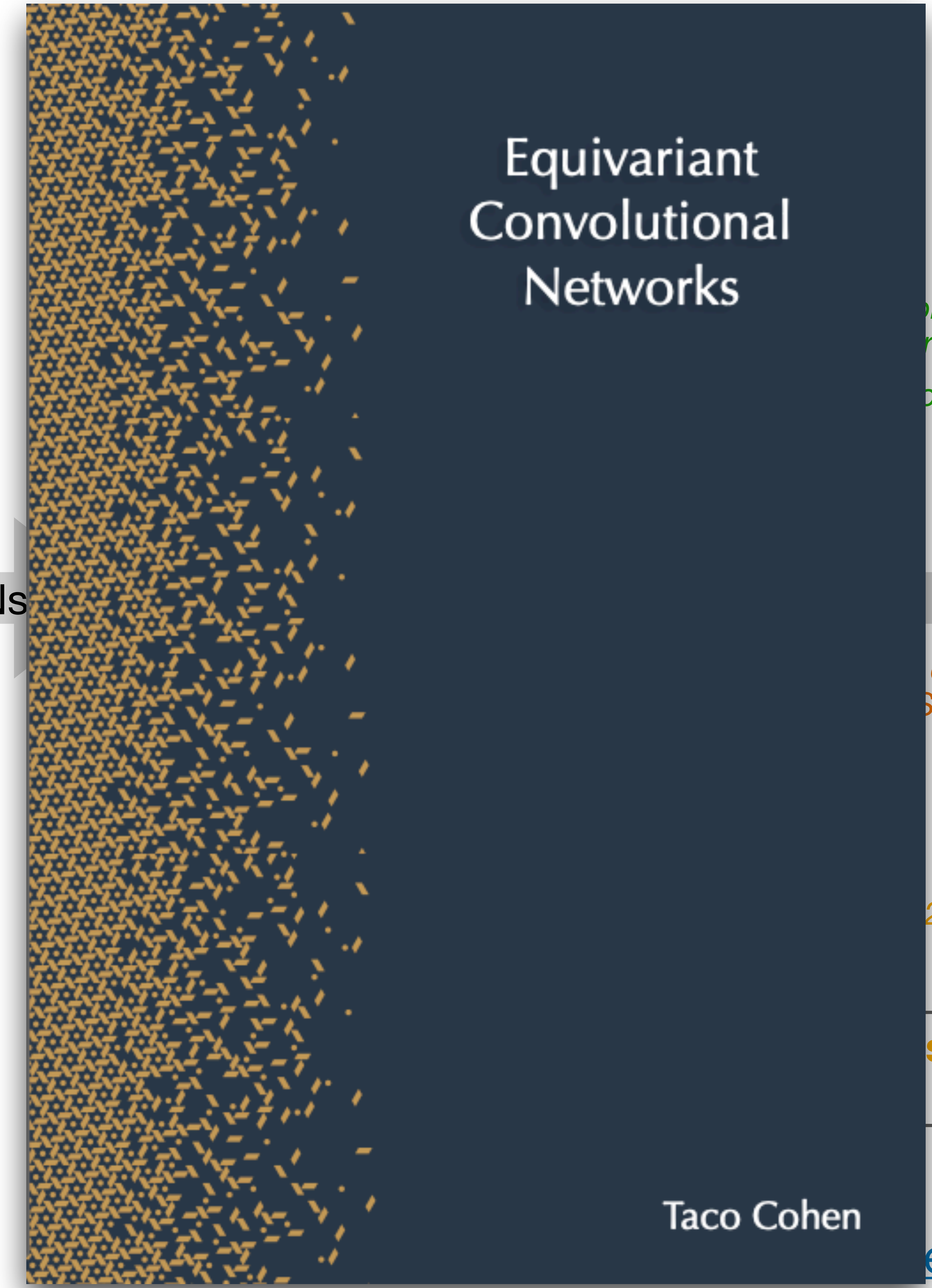
...e3nn/e3nn

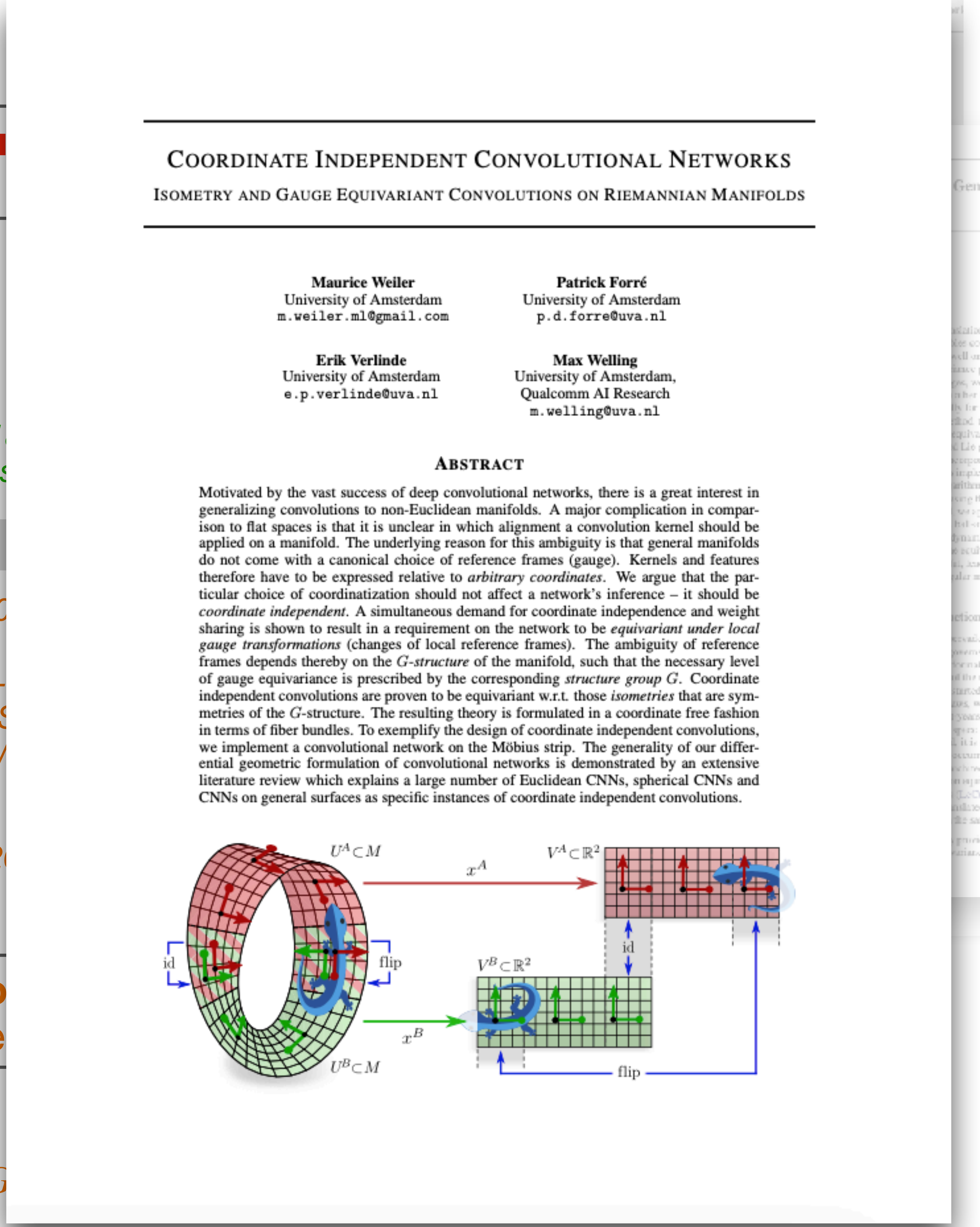
https://pure.uva.nl/ws/files/60770359/Thesis.pdf

https://quva-lab.github.io/escnn/

From plain NNs to Gauge-equivariant CNNs

https://github.com/Chen-Cai-OSU/awesome-equivariant-network





COORDINATE INDEPENDENT CONVOLUTIONAL NETWORKS

ISOMETRY AND GAUGE EQUIVARIANT CONVOLUTIONS ON RIEMANNIAN MANIFOLDS

Maurice Weiler
University of Amsterdam
m.weiler.ml@gmail.com

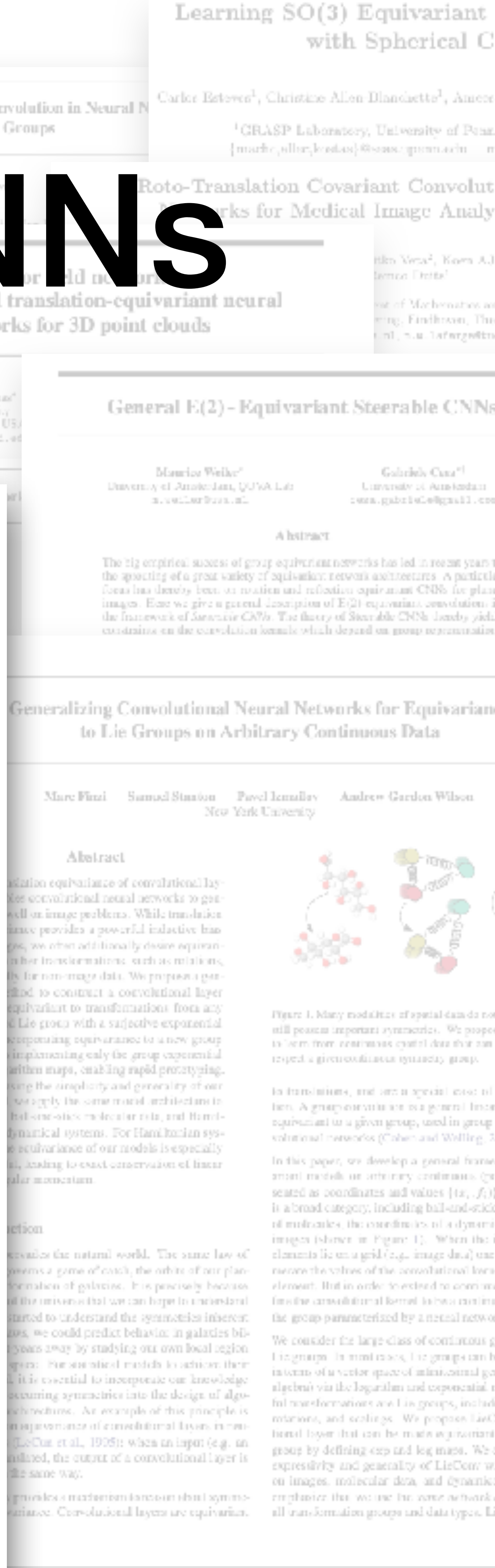
Patrick Forré
University of Amsterdam
p.d.forre@uva.nl

Erik Verlinde
University of Amsterdam
e.p.verlinde@uva.nl

Max Welling
University of Amsterdam,
Qualcomm AI Research
m.welling@uva.nl

ABSTRACT

Motivated by the vast success of deep convolutional networks, there is a great interest in generalizing convolutions to non-Euclidean manifolds. A major complication in comparison to flat spaces is that it is unclear in which alignment a convolution kernel should be applied on a manifold. The underlying reason for this ambiguity is that general manifolds do not come with a canonical choice of reference frames (gauge). Kernels and features therefore have to be expressed relative to *arbitrary coordinates*. We argue that the particular choice of coordinatization should not affect a network's inference – it should be *coordinate independent*. A simultaneous demand for coordinate independence and weight sharing is shown to result in a requirement on the network to be *equivariant under local gauge transformations* (changes of local reference frames). The ambiguity of reference frames depends thereby on the *G-structure* of the manifold, such that the necessary level of gauge equivariance is prescribed by the corresponding *structure group G*. Coordinate independent convolutions are proven to be equivariant w.r.t. those *isometries* that are symmetries of the *G*-structure. The resulting theory is formulated in a coordinate free fashion in terms of fiber bundles. To exemplify the design of coordinate independent convolutions, we implement a convolutional network on the Möbius strip. The generality of our differential geometric formulation of convolutional networks is demonstrated by an extensive literature review which explains a large number of Euclidean CNNs, spherical CNNs and CNNs on general surfaces as specific instances of coordinate independent convolutions.



https://pure.uva.nl/ws/files/60770359/Thesis.pdf

https://quva-lab.github.io/escnn/