Learning Group Structured Representations

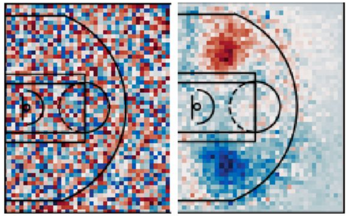
CS7180 Geometric DL

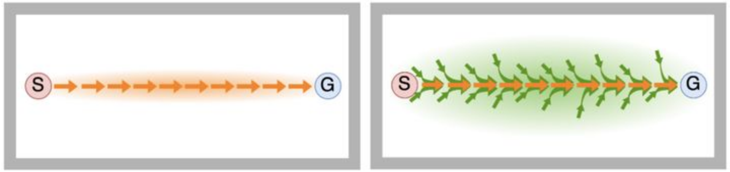
John Park

March 23 2023

Bio

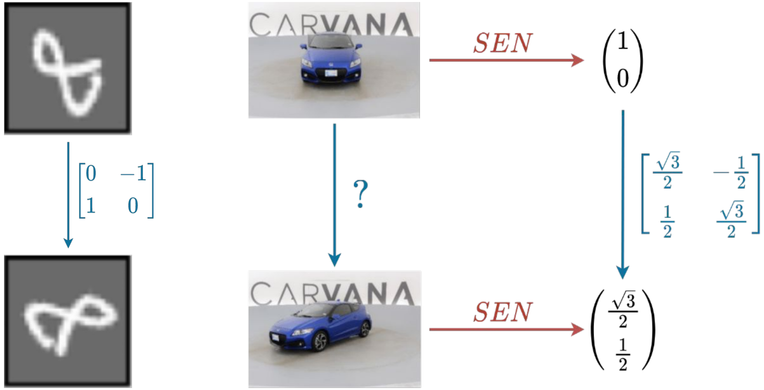
Reinforcement learning and symmetry





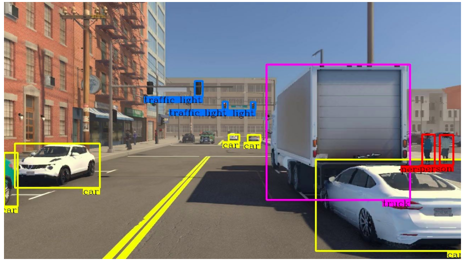
(c) BC (d) BMIL

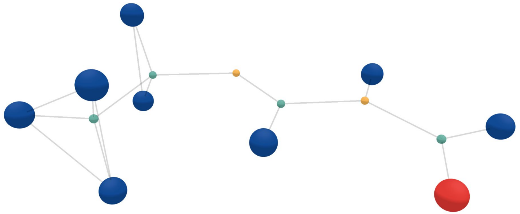
Strong inductive biases to speed up learning and improve generalization

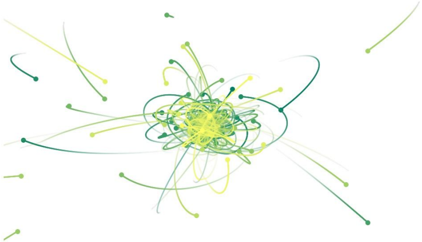


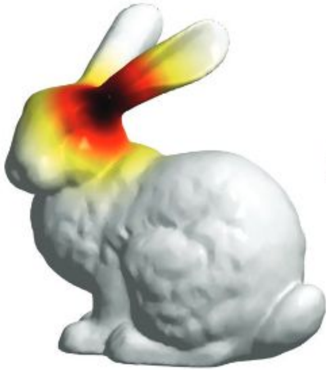
Symmetry

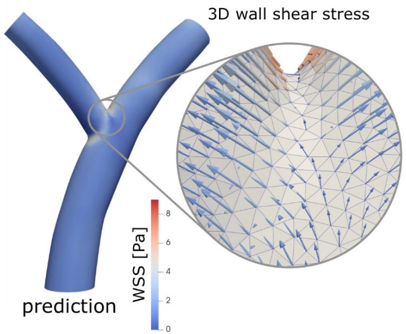
Important geometric prior







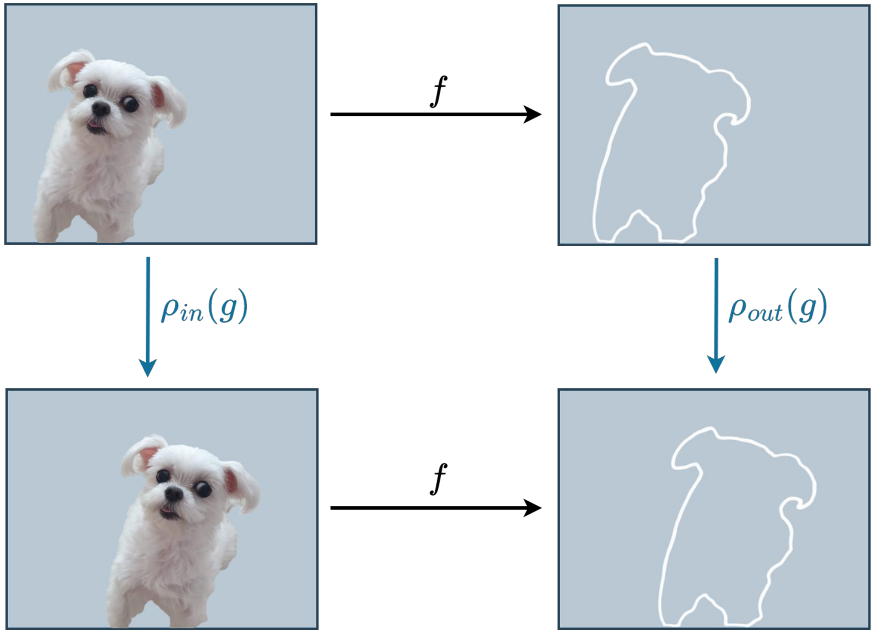




References

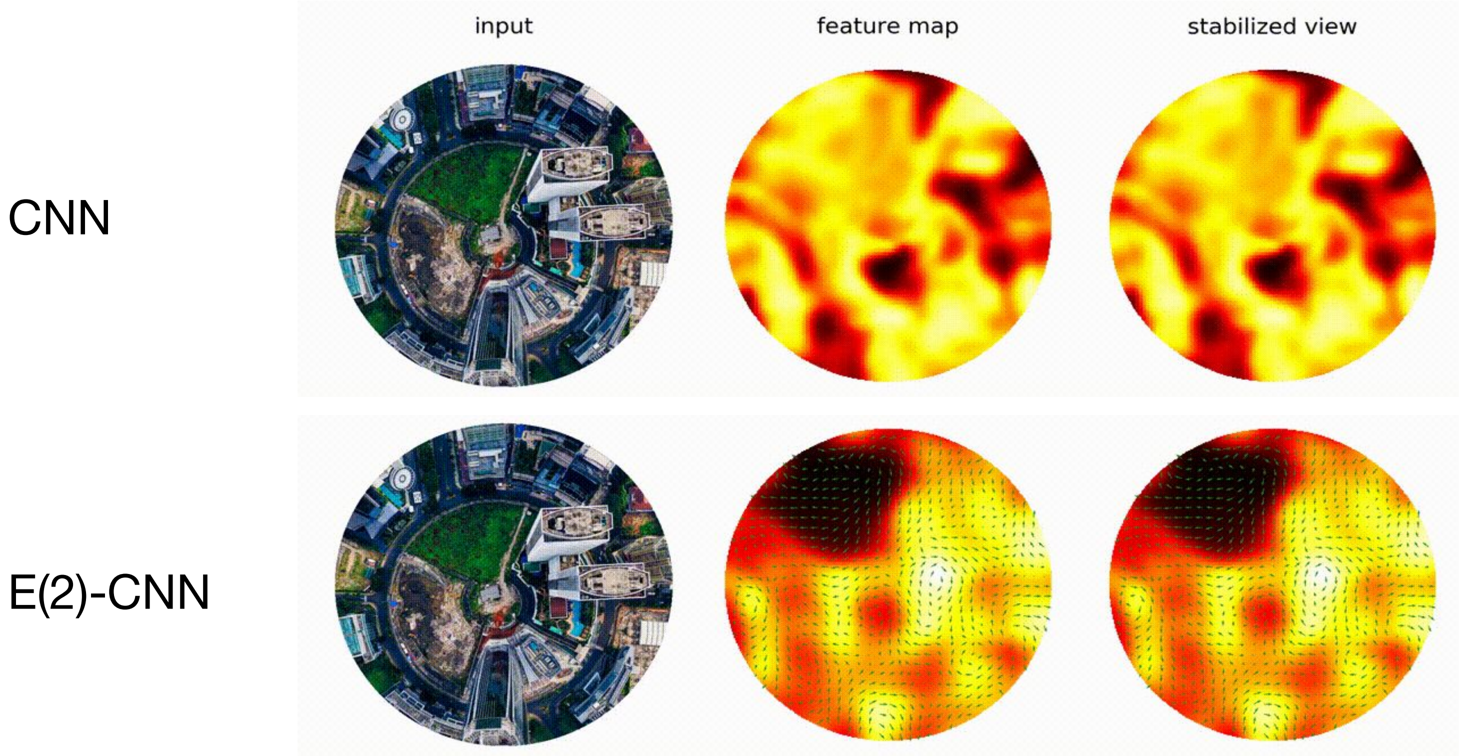
1. https://developer.nvidia.com/blog/deploying-a-scalable-object-detection-pipeline-the-inferencing-process-part-2/
2. https://robdhess.github.io/Steerable-E3-GNN/
3. Monti, Federico, et al. "Geometric deep learning on graphs and manifolds using mixture model cnns." CVPR 2017.
4. Bronstein, Michael M., et al. "Geometric deep learning: going beyond euclidean data." IEEE Signal Processing Magazine 34.4 (2017): 18-42.
5. https://github.com/sukjulian/coronary-mesh-convolution

Equivariance



$$f(\rho_{in}(g)x) = \rho_{out}(g)f(x)$$

Equivariance



References

1. https://github.com/QUVA-Lab/e2cnn

Limitations

$$f(\rho_{in}(g)x) = \rho_{out}(g)f(x)$$

Symmetry group needs to be known

Limitations

$$f(\rho_{in}(g)x) = \rho_{out}(g)f(x)$$

Symmetry group needs to be known

Group action on input and output spaces must be known

Limitations

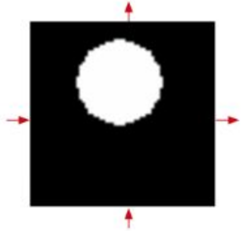
$$f(\rho_{in}(g)x) = \rho_{out}(g)f(x)$$

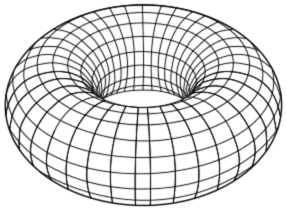
Symmetry group needs to be known

Group action on input and output spaces must be known

Equivariant model tied to group and action

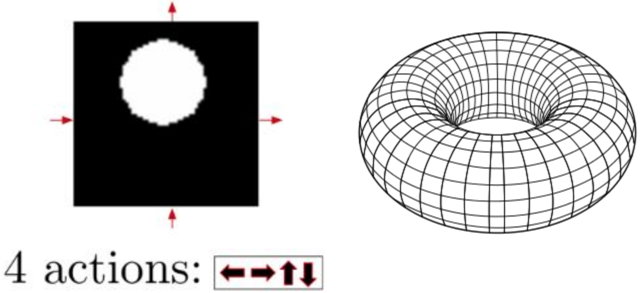
Transformations in latent space

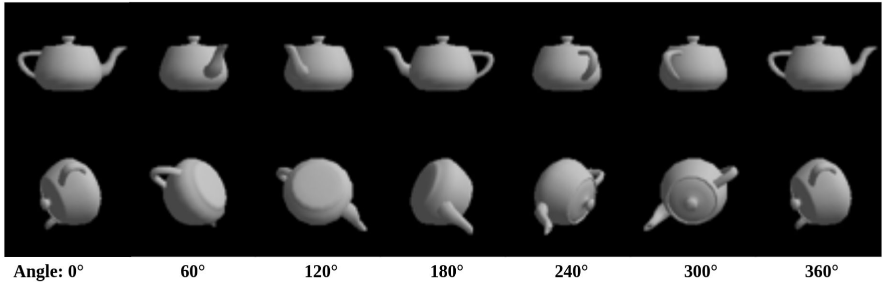




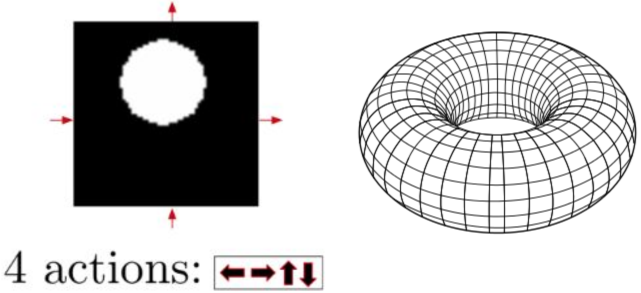
4 actions: ←→↑↓

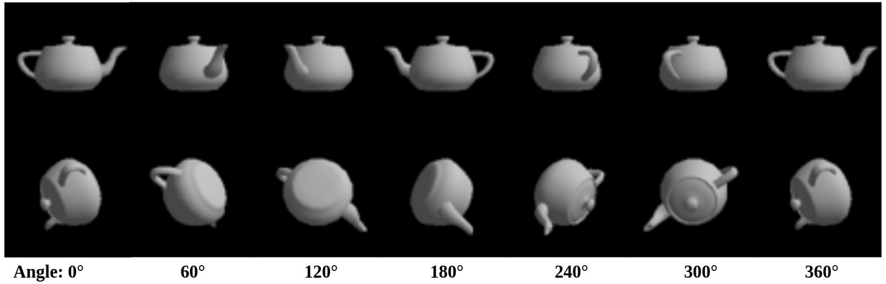
Transformations in latent space





Transformations in latent space





The mug is on the right of the robot gripper

↓

The mug is in front of the robot gripper

Learning symmetry from data

Learning the Irreducible Representations of Commutative Lie Groups

Taco Cohen T.S.COHEN@UVA.NL
Max Welling M.WELLING@UVA.NL
Machine Learning Group, University of Amsterdam

Abstract

We present a new probabilistic model of compact commutative Lie groups that produces invariant-equivariant and disentangled representations of data. To define the notion of disentangling, we borrow a fundamental principle from physics that is used to derive the elementary particles of a system from its symmetries. Our model employs a newfound Bayesian conjugacy relation that enables fully tractable probabilistic inference over compact commutative Lie groups – a class that includes the groups that describe the rotation and cyclic translation of images. We train the model on pairs of transformed image patches, and show that the learned invariant representation is highly effective for classification.

1. Introduction

Recently, the field of deep learning has produced some remarkable breakthroughs. The hallmark of the deep learning approach is to learn multiple layers of *representation* of data, and much work has gone into the develop-

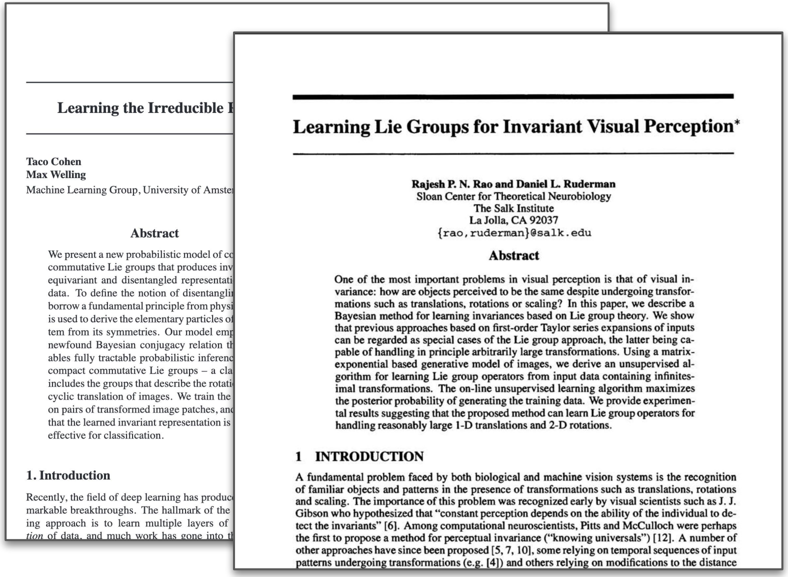
for meaning for future work.

What do we mean, intuitively, when we speak of invariance and disentangling? A disentangled representation is one that explicitly represents the distinct factors of variation in the data. For example, visual data (i.e. pixels) can be thought of as a composition of object identity, position and pose, lighting conditions, etc. Once disentangling is achieved, invariance follows easily: to build a representation that is invariant to the transformation of a factor of variation (e.g. object position) that is considered a nuisance for a particular task (e.g. object classification), one can simply ignore the units in the representation that encode the nuisance factor.

To get a mathematical handle on the concept of disentangling, we borrow a fundamental principle from physics, which we refer to as Weyl's principle, following Kanatani (1990). In physics, this idea is used to tease apart (i.e. disentangle) the *elementary particles* of a physical system from mere measurement values that have no inherent physical significance. We apply this principle to the area of vision, for after all, pixels are nothing but physical measurements.

Weyl's principle presupposes a symmetry group that acts

Learning symmetry from data



Learning the Irreducible R

Taco Cohen
Max Welling
Machine Learning Group, University of Amster

Abstract

We present a new probabilistic model of co
commutative Lie groups that produces inv
equivariant and disentangled representati
data. To define the notion of disentanglin
borrow a fundamental principle from physi
is used to derive the elementary particles of
tem from its symmetries. Our model emp
newfound Bayesian conjugacy relation th
ables fully tractable probabilistic inferenc
compact commutative Lie groups – a cla
includes the groups that describe the rotati
cyclic translation of images. We train the
on pairs of transformed image patches, and
that the learned invariant representation is
effective for classification.

1. Introduction

Recently, the field of deep learning has produc
markable breakthroughs. The hallmark of the
ing approach is to learn multiple layers of
tion of data, and much work has gone into th

Learning Lie Groups for Invariant Visual Perception*

Rajesh P. N. Rao and Daniel L. Ruderman
Sloan Center for Theoretical Neurobiology
The Salk Institute
La Jolla, CA 92037
{rao,ruderman}@salk.edu

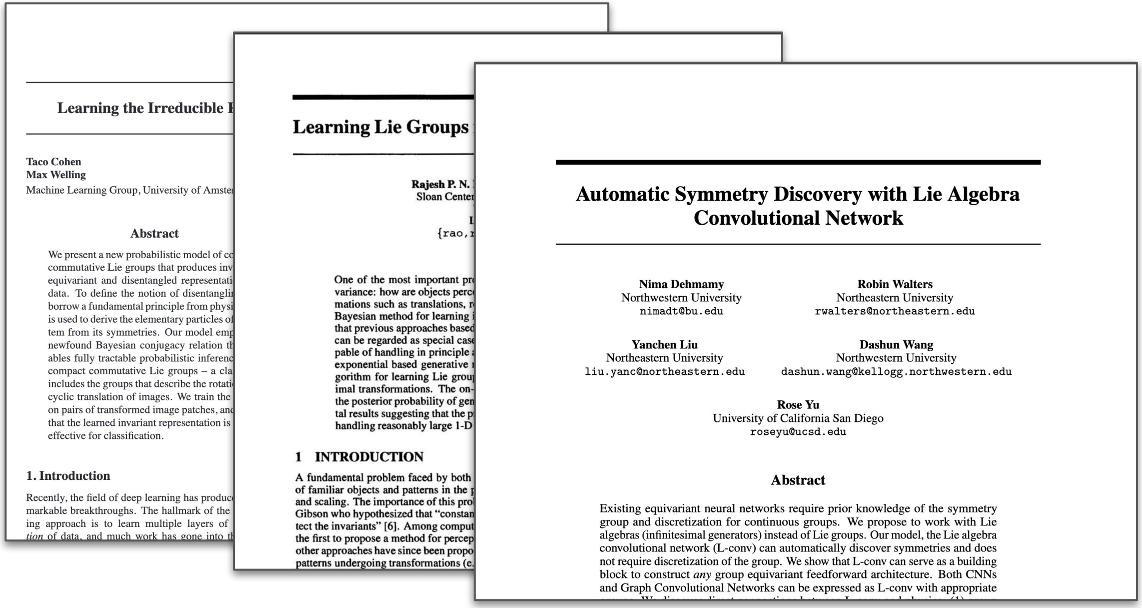
Abstract

One of the most important problems in visual perception is that of visual invariance: how are objects perceived to be the same despite undergoing transformations such as translations, rotations or scaling? In this paper, we describe a Bayesian method for learning invariances based on Lie group theory. We show that previous approaches based on first-order Taylor series expansions of inputs can be regarded as special cases of the Lie group approach, the latter being capable of handling in principle arbitrarily large transformations. Using a matrix-exponential based generative model of images, we derive an unsupervised algorithm for learning Lie group operators from input data containing infinitesimal transformations. The on-line unsupervised learning algorithm maximizes the posterior probability of generating the training data. We provide experimental results suggesting that the proposed method can learn Lie group operators for handling reasonably large 1-D translations and 2-D rotations.

1 INTRODUCTION

A fundamental problem faced by both biological and machine vision systems is the recognition of familiar objects and patterns in the presence of transformations such as translations, rotations and scaling. The importance of this problem was recognized early by visual scientists such as J. J. Gibson who hypothesized that "constant perception depends on the ability of the individual to detect the invariants" [6]. Among computational neuroscientists, Pitts and McCulloch were perhaps the first to propose a method for perceptual invariance ("knowing universals") [12]. A number of other approaches have since been proposed [5, 7, 10], some relying on temporal sequences of input patterns undergoing transformations (e.g. [4]) and others relying on modifications to the distance

Learning symmetry from data



Automatic Symmetry Discovery with Lie Algebra Convolutional Network

Nima Dehmamy
Northwestern University
nimadt@bu.edu

Robin Walters
Northeastern University
rwalters@northeastern.edu

Yanchen Liu
Northeastern University
liu.yanc@northeastern.edu

Dashun Wang
Northwestern University
dashun.wang@kellogg.northwestern.edu

Rose Yu
University of California San Diego
roseyu@ucsd.edu

Abstract

Existing equivariant neural networks require prior knowledge of the symmetry group and discretization for continuous groups. We propose to work with Lie algebras (infinitesimal generators) instead of Lie groups. Our model, the Lie algebra convolutional network (L-conv) can automatically discover symmetries and does not require discretization of the group. We show that L-conv can serve as a building block to construct *any* group equivariant feedforward architecture. Both CNNs and Graph Convolutional Networks can be expressed as L-conv with appropriate

Learning symmetry from data

Published as a conference paper at ICLR 2021

META-LEARNING SYMMETRIES BY REPARAMETERIZATION

Allan Zhou, Tom Knowles, Chelsea Finn
Dept of Computer Science, Stanford University
{ayz,tknowles,cbfinn}@stanford.edu

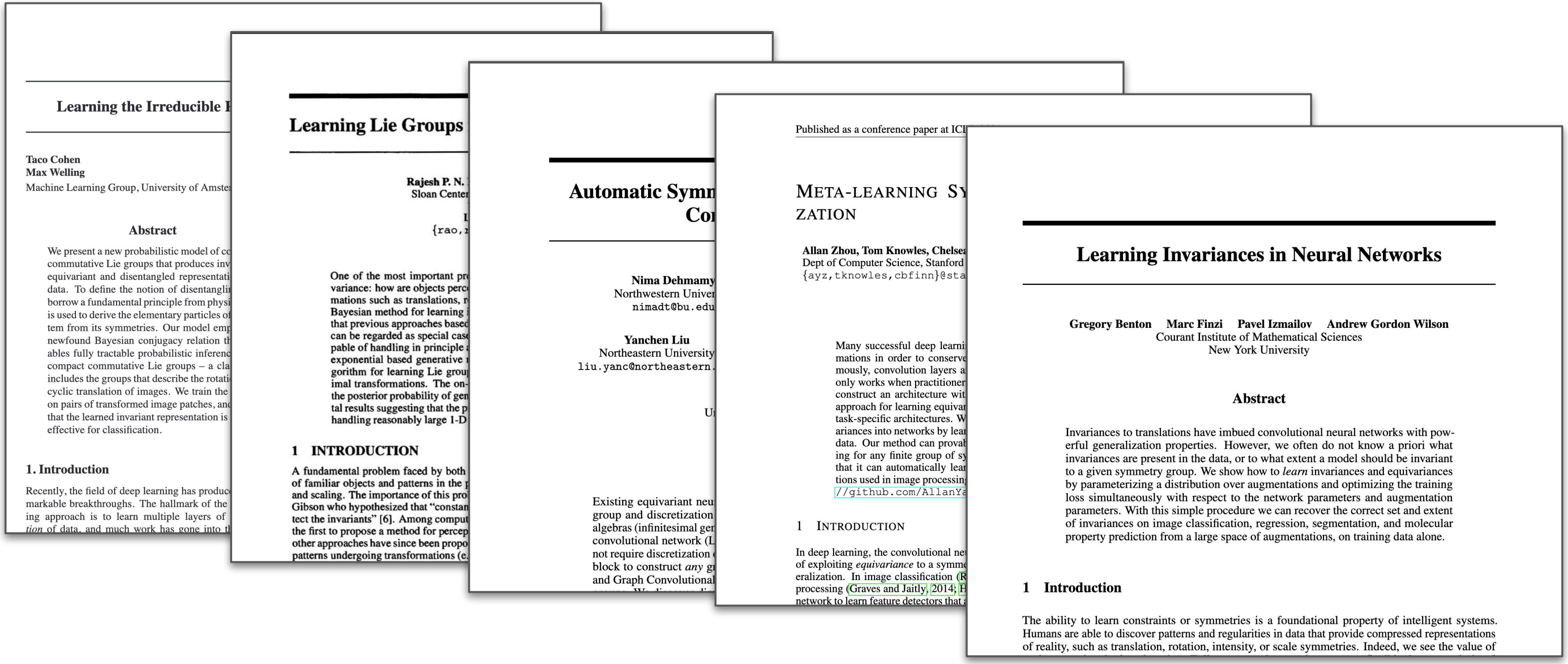
ABSTRACT

Many successful deep learning architectures are equivariant to certain transformations in order to conserve parameters and improve generalization: most famously, convolution layers are equivariant to shifts of the input. This approach only works when practitioners know the symmetries of the task and can manually construct an architecture with the corresponding equivariances. Our goal is an approach for learning equivariances from data, without needing to design custom task-specific architectures. We present a method for learning and encoding equivariances into networks by learning corresponding parameter sharing patterns from data. Our method can provably represent equivariance-inducing parameter sharing for any finite group of symmetry transformations. Our experiments suggest that it can automatically learn to encode equivariances to common transformations used in image processing tasks. We provide our experiment code at https://github.com/AllanYangZhou/metalearning-symmetries.

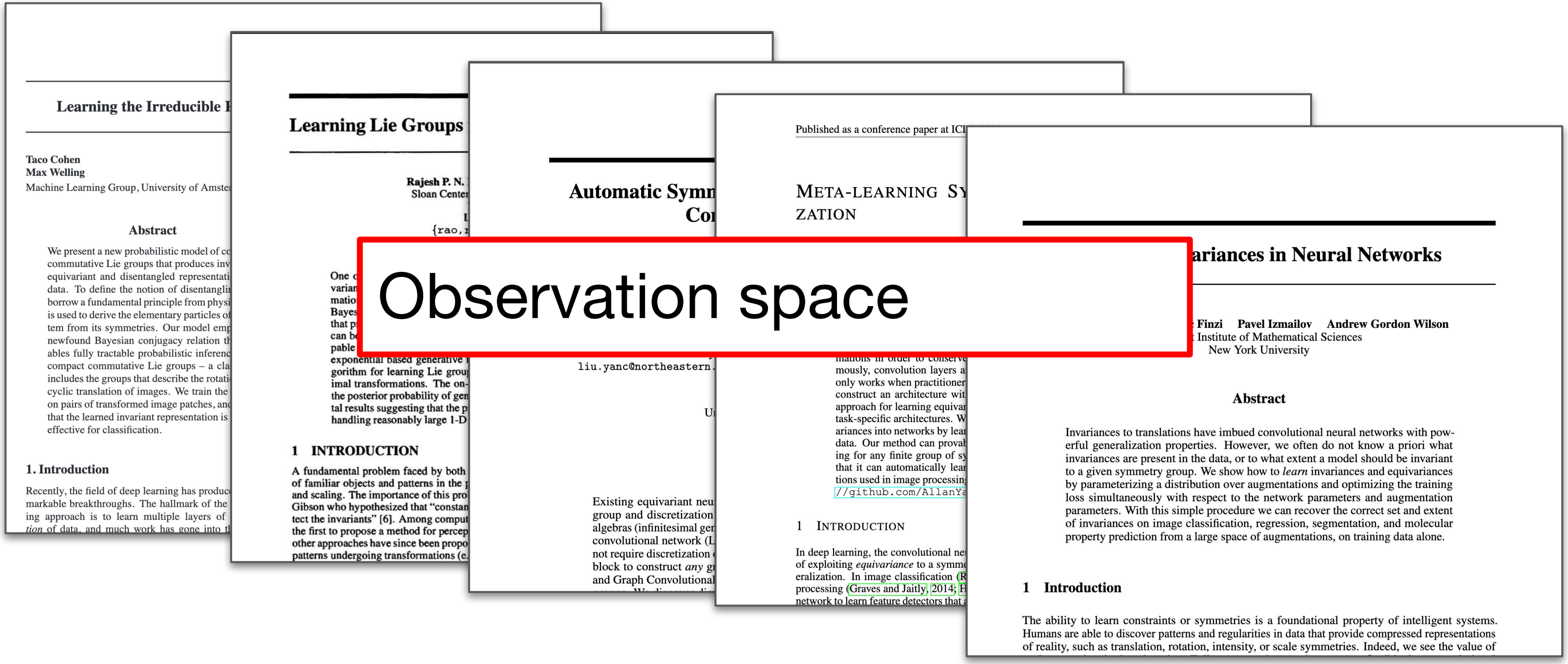
1 INTRODUCTION

In deep learning, the convolutional neural network (CNN) (LeCun et al., 1998) is a prime example of exploiting *equivariance* to a symmetry transformation to conserve parameters and improve generalization. In image classification (Russakovsky et al., 2015; Krizhevsky et al., 2012) and audio processing (Graves and Jaitly, 2014; Hannun et al., 2014) tasks, we may expect the layers of a deep network to learn feature detectors that are translation equivariant: if we translate the input, the output

Learning symmetry from data



Learning symmetry from data



Outline

Group-structured Representations (Higgins et al. 2018)

Learning disentangled representations (Quessard et al. 2020)

Learning symmetric representations (Park et al. 2022)

Representations

What is a good representation?

- Makes learning task easier
- Preserve underlying structure/information of data

References

1. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35(8), 1798-1828.
2. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.
3. Locatello, F., Bauer, S., Lucic, M., Raetsch, G., Gelly, S., Schölkopf, B., & Bachem, O. (2019, May). Challenging common assumptions in the unsupervised learning of disentangled representations. ICML
4. Higgins, I., Amos, D., Pfau, D., Racaniere, S., Matthey, L., Rezende, D., & Lerchner, A. (2018). Towards a definition of disentangled representations. arXiv preprint arXiv:1812.02230.

Representations

What is a good representation?

- Makes learning task easier
- Preserve underlying structure/information of data

Properties

- Lower-dimensional → faster learning
- Useful for other downstream tasks
- Disentangled (separate factors of variation)

References

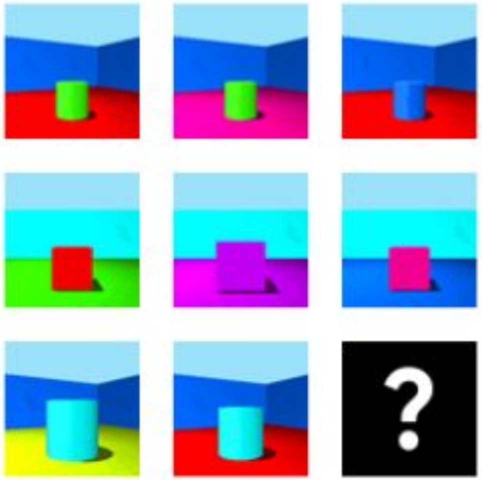
1. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35(8), 1798-1828.
2. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.
3. Locatello, F., Bauer, S., Lucic, M., Raetsch, G., Gelly, S., Schölkopf, B., & Bachem, O. (2019, May). Challenging common assumptions in the unsupervised learning of disentangled representations. ICML
4. Higgins, I., Amos, D., Pfau, D., Racaniere, S., Matthey, L., Rezende, D., & Lerchner, A. (2018). Towards a definition of disentangled representations. arXiv preprint arXiv:1812.02230.

Disentangled Representations

Good representations should disentangle the **factors of the data generating process**

- Color
- Shape
- Background

No agreement on exact definition



References

1. Van Steenkiste, S., Locatello, F., Schmidhuber, J., & Bachem, O. (2019). Are disentangled representations helpful for abstract visual reasoning?. NeurIPS

Disentangled representations

Towards a Definition of Disentangled Representations

Irina Higgins*, David Amos*, David Pfau, Sebastien Racaniere, Loic Matthey, Danilo Rezende, Alexander Lerchner
DeepMind
{irinah,davidamos,pfau,sracaniere, lmatthey,danilor,lerchner}@google.com

December 7, 2018

How can intelligent agents solve a diverse set of tasks in a data-efficient manner? The disentangled representation learning approach posits that such an agent would benefit from separating out (disentangling) the underlying structure of the world into disjoint parts of its representation. However, there is no generally agreed-upon definition of disentangling, not least because it is unclear how to formalise the notion of world structure beyond toy datasets with a known ground truth generative process. Here we propose that a principled solution to characterising disentangled representations can be found

Connects disentanglement with symmetry inspired by physics

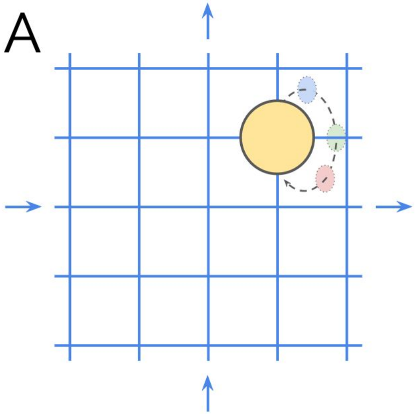
Shift focus from objects to transformations

Observe how an object transforms to learn disentangled representation

Disentangled Representation

A vector representation is called a ***disentangled representation*** *with respect to a particular decomposition of a symmetry group into subgroups, if it decomposes into independent subspaces, where each subspace is affected by the action of a single subgroup, and the actions of all other subgroups leave the subspace unaffected.*

A



$$G = G_x \times G_y \times G_c$$

$$Z = Z_x \times Z_y \times Z_c$$

Paper 1

Learning Disentangled Representations and Group Structure of Dynamical Environments

Robin Quessard[1,2*] **Thomas D. Barrett**[3*] **William R. Clements**[1*]
[1]Indust.ai, Paris, France
[2]École Normale Supérieure, Paris, France
[3]University of Oxford, Oxford, UK
robin.quessard@ens.fr, thomas.barrett@physics.ox.ac.uk, william.clements@indust.ai

Abstract

Learning disentangled representations is a key step towards effectively discovering and modelling the underlying structure of environments. In the natural sciences, physics has found great success by describing the universe in terms of symmetry preserving transformations. Inspired by this formalism, we propose a framework, built upon the theory of group representation, for learning representations of a dynamical environment structured around the transformations that generate its evolution. Experimentally, we learn the structure of explicitly symmetric environments without supervision from observational data generated by sequential interactions. We further introduce an intuitive disentanglement regularisation to ensure the interpretability of the learnt representations. We show that our method enables accurate long-horizon predictions, and demonstrate a correlation between the quality of predictions and disentanglement in the latent space.

Practical method to learn representations using definition from Higgins et al. 2018.

Focuses on SO(n)

Considers a trajectory of objects and transformations

Method

Dataset: $(o_1, a_1, o_2 \ldots, a_T, o_T)$

Goal: Learn encoder $f_\phi : O \to V$, $a_i \mapsto g_i \in GL(V)$, decoder $d_\psi : V \mapsto O$

$$\hat{o}_{i+m}(\phi, \psi, \theta) = d_\psi(g_{i+m}(\theta).g_{i+m-1}(\theta)....g_{i+1}(\theta).f_\phi(o_i))$$

Method

Parameterize g as product of rotations

$$g(\theta_{1,2}, \theta_{1,3} .., \theta_{n-1,n}) = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} R_{i,j}(\theta_{i,j})$$

$$R = R_z(\alpha)\, R_y(\beta)\, R_x(\gamma) = \overset{\text{yaw}}{\begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}} \overset{\text{pitch}}{\begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}} \overset{\text{roll}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{bmatrix}}$$

Method

$$\hat{o}_{i+m}(\phi, \psi, \theta) = d_\psi(g_{i+m}(\theta).g_{i+m-1}(\theta)....g_{i+1}(\theta).f_\phi(o_i))$$

Learn parameters ϕ , ψ , θ

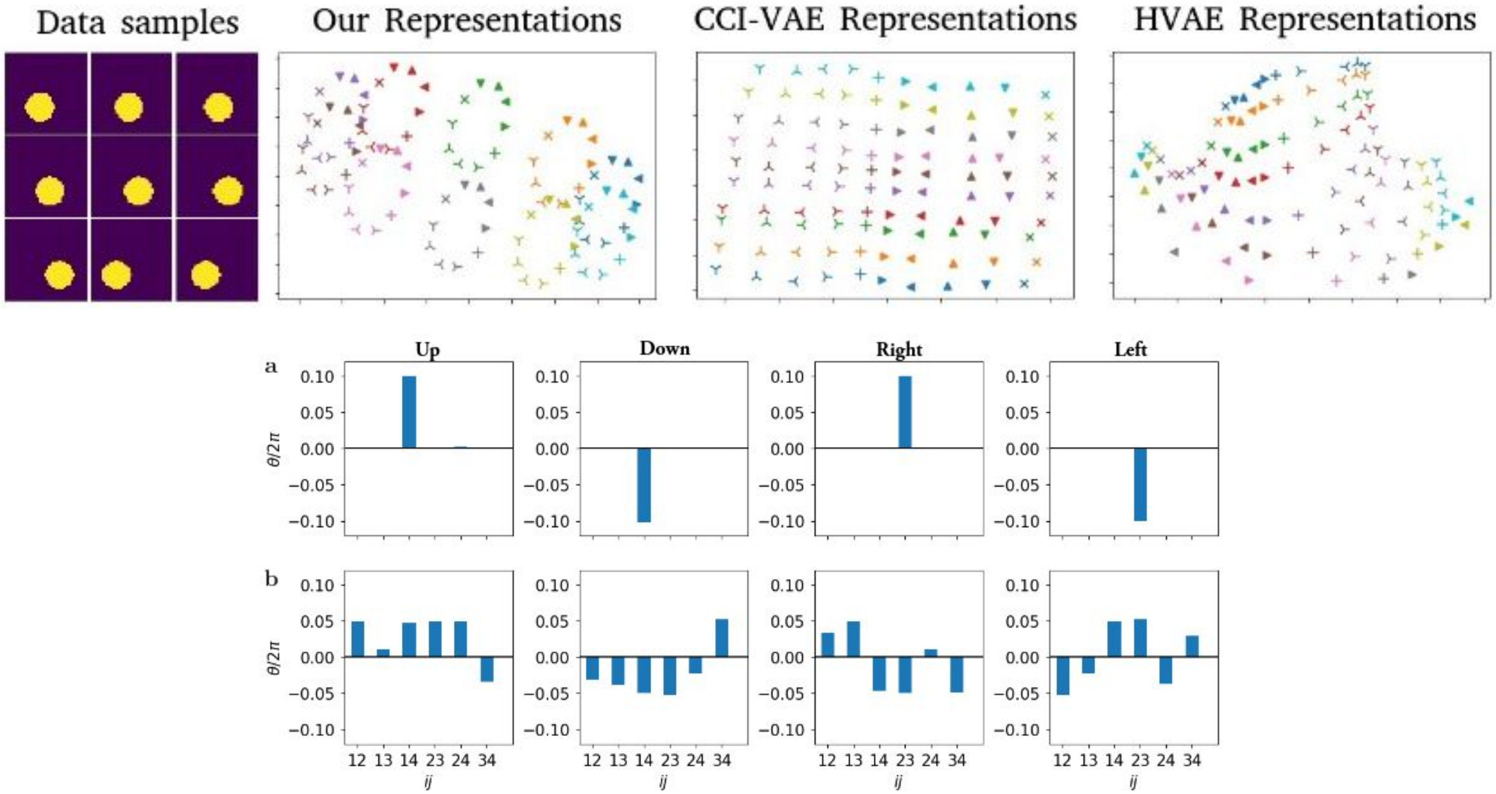
$$\mathcal{L}_{total} = \mathcal{L}_{rec} + \lambda \mathcal{L}_{ent}$$

$$\mathcal{L}_{rec} = l(o_{i+1}, \hat{o}_{i+1})$$

$$\mathcal{L}_{\text{ent}}(\theta) = \sum_a \sum_{(i,j) \neq (\alpha,\beta)} |\theta^a_{i,j}|^2 \qquad \text{with} \qquad \theta^a_{\alpha,\beta} = \max_{i,j}(|\theta^a_{i,j}|)$$

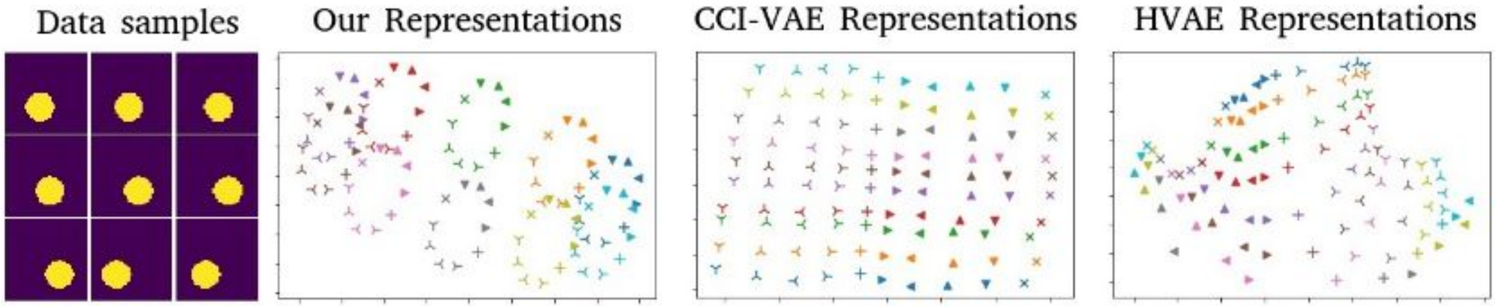
Minimize entanglement (promotes sparsity)

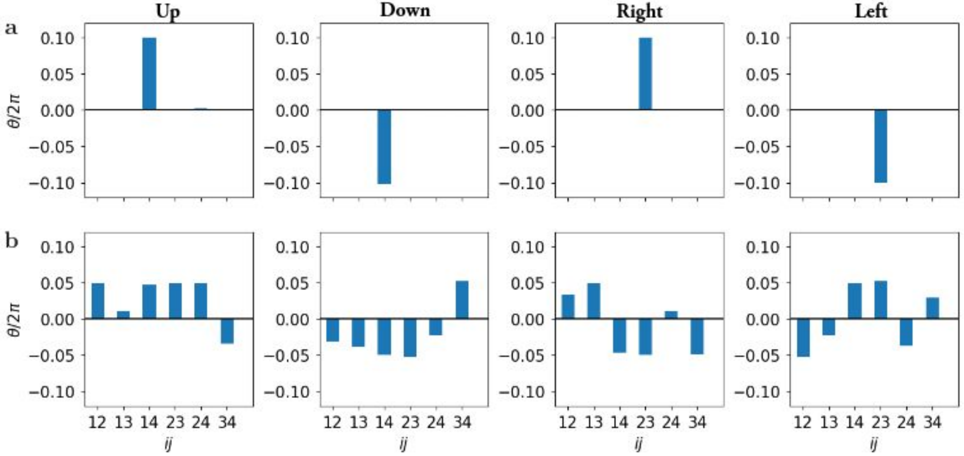
Experiments (Flatland)

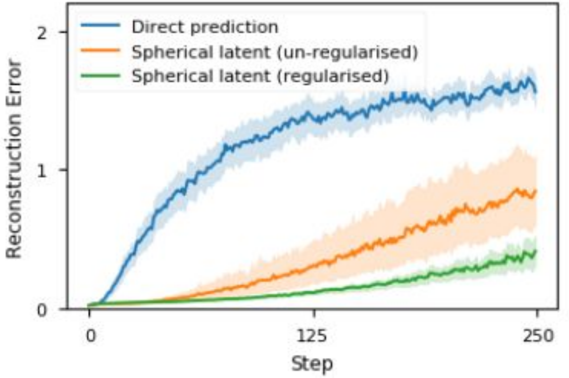


Experiments (Flatland)

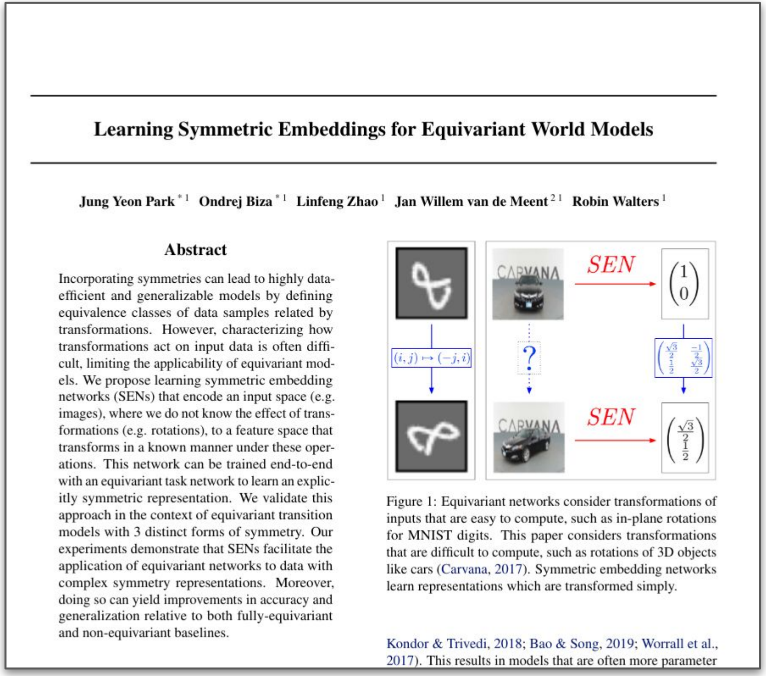
$$G = C_p \times C_p$$







Paper 2



Learning Symmetric Embeddings for Equivariant World Models

Jung Yeon Park [*1] Ondrej Biza [*1] Linfeng Zhao [1] Jan Willem van de Meent [2 1] Robin Walters [1]

Abstract

Incorporating symmetries can lead to highly data-efficient and generalizable models by defining equivalence classes of data samples related by transformations. However, characterizing how transformations act on input data is often difficult, limiting the applicability of equivariant models. We propose learning symmetric embedding networks (SENs) that encode an input space (e.g. images), where we do not know the effect of transformations (e.g. rotations), to a feature space that transforms in a known manner under these operations. This network can be trained end-to-end with an equivariant task network to learn an explicitly symmetric representation. We validate this approach in the context of equivariant transition models with 3 distinct forms of symmetry. Our experiments demonstrate that SENs facilitate the application of equivariant networks to data with complex symmetry representations. Moreover, doing so can yield improvements in accuracy and generalization relative to both fully-equivariant and non-equivariant baselines.

Figure 1: Equivariant networks consider transformations of inputs that are easy to compute, such as in-plane rotations for MNIST digits. This paper considers transformations that are difficult to compute, such as rotations of 3D objects like cars (Carvana, 2017). Symmetric embedding networks learn representations which are transformed simply.

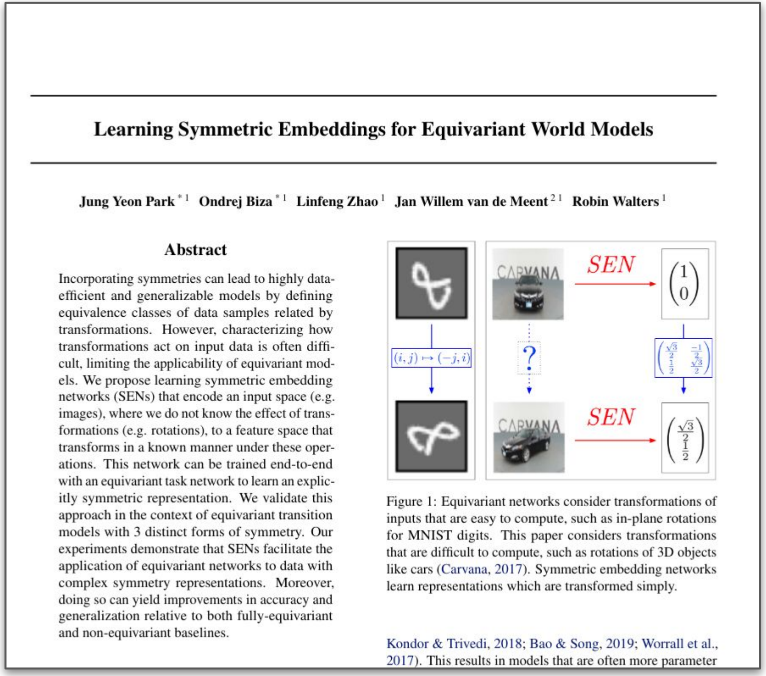
Kondor & Trivedi, 2018; Bao & Song, 2019; Worrall et al., 2017). This results in models that are often more parameter

Ignores disentanglement

Arbitrary groups and domains

Actions do not need to be tied to transformations

Paper 2



Learning Symmetric Embeddings for Equivariant World Models

Jung Yeon Park [*1] Ondrej Biza [*1] Linfeng Zhao [1] Jan Willem van de Meent [2 1] Robin Walters [1]

Abstract

Incorporating symmetries can lead to highly data-efficient and generalizable models by defining equivalence classes of data samples related by transformations. However, characterizing how transformations act on input data is often difficult, limiting the applicability of equivariant models. We propose learning symmetric embedding networks (SENs) that encode an input space (e.g. images), where we do not know the effect of transformations (e.g. rotations), to a feature space that transforms in a known manner under these operations. This network can be trained end-to-end with an equivariant task network to learn an explicitly symmetric representation. We validate this approach in the context of equivariant transition models with 3 distinct forms of symmetry. Our experiments demonstrate that SENs facilitate the application of equivariant networks to data with complex symmetry representations. Moreover, doing so can yield improvements in accuracy and generalization relative to both fully-equivariant and non-equivariant baselines.

Figure 1: Equivariant networks consider transformations of inputs that are easy to compute, such as in-plane rotations for MNIST digits. This paper considers transformations that are difficult to compute, such as rotations of 3D objects like cars (Carvana, 2017). Symmetric embedding networks learn representations which are transformed simply.

Kondor & Trivedi, 2018; Bao & Song, 2019; Worrall et al., 2017). This results in models that are often more parameter

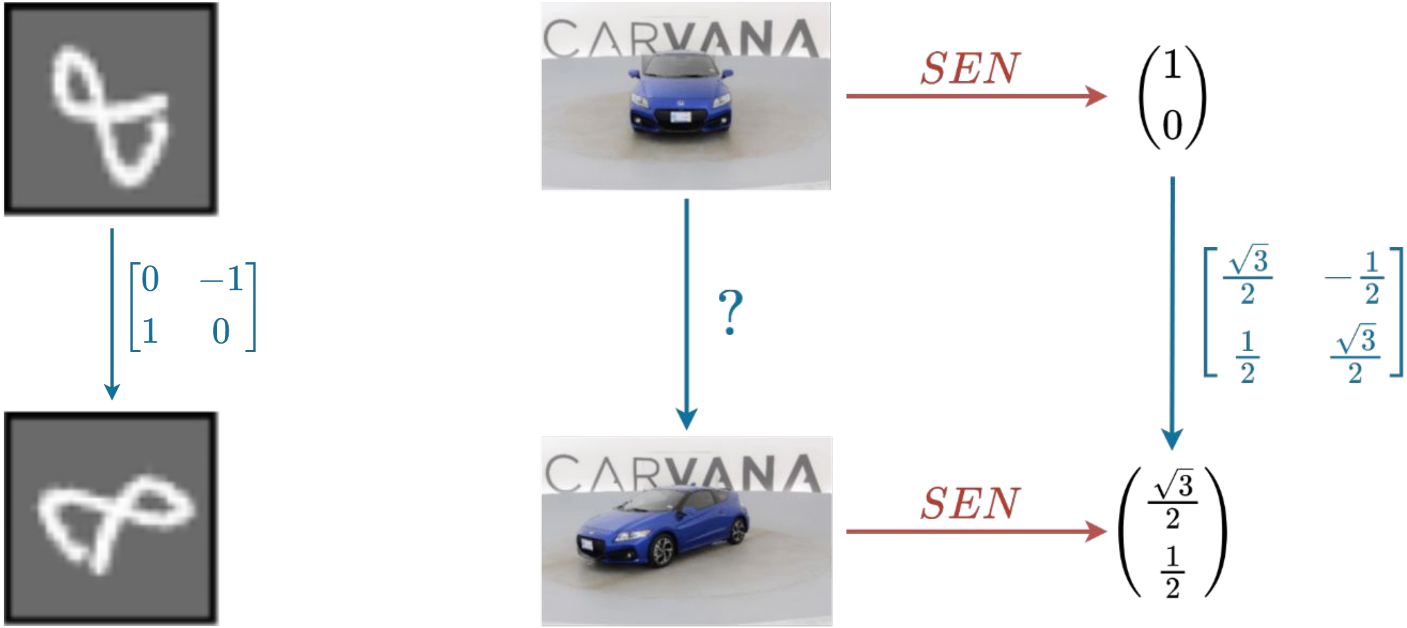
Ignores disentanglement

Arbitrary groups and domains

Actions do not need to be tied to transformations

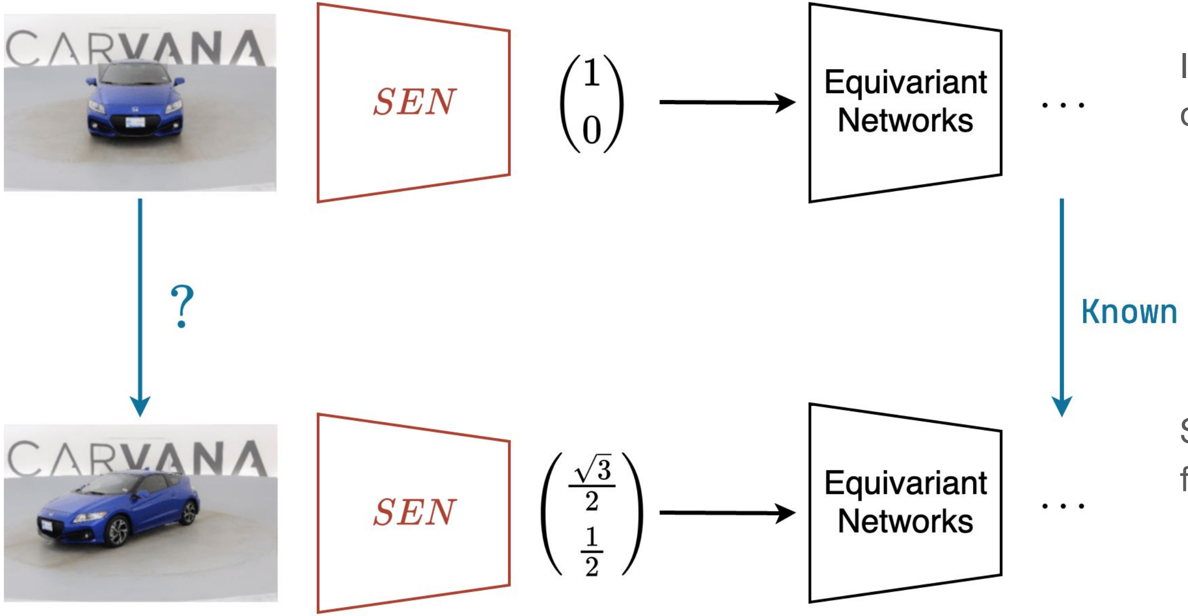
Symmetric Embedding Network (SEN)

Difficult to characterize how symmetry acts on input data, i.e. group action is unknown



Symmetric Embedding Network (SEN)

Key idea: Pair SEN with downstream equivariant networks and train end-to-end

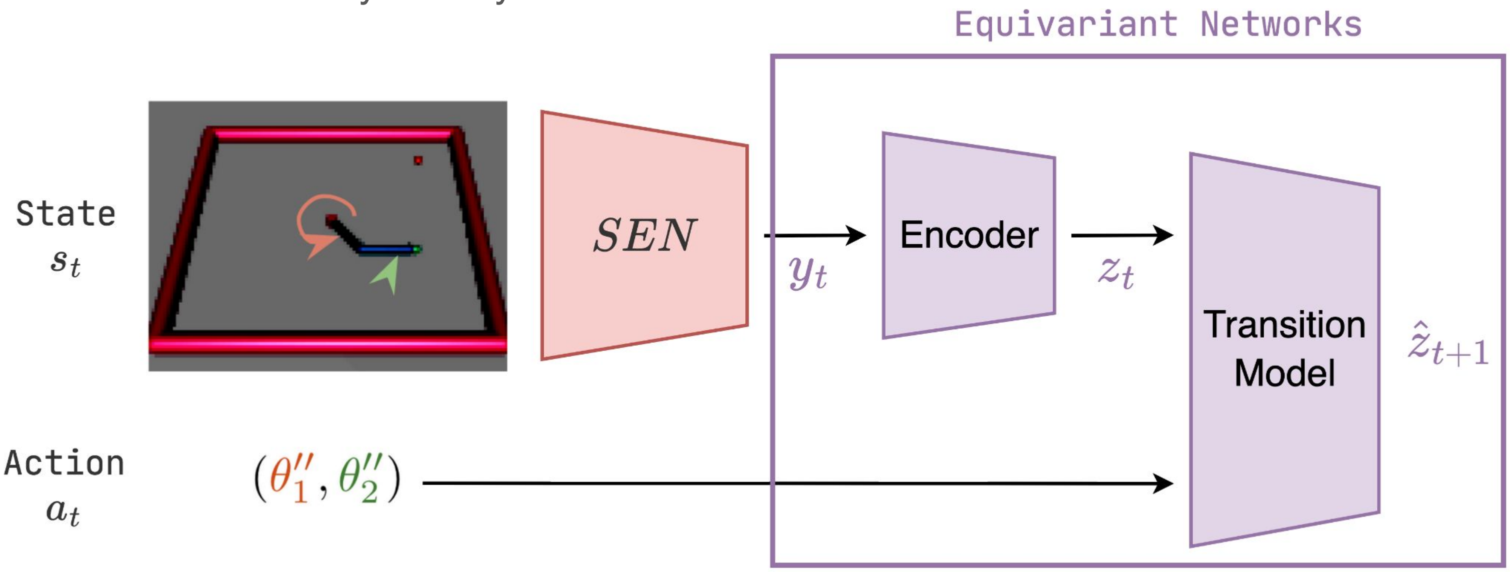


Indirect supervision from downstream equivariant networks

SEN learns to map from image space to feature space with known group action

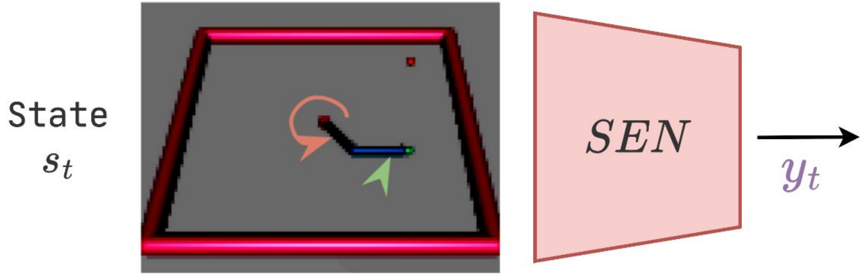
Meta-Architecture

Learn world models with symmetry



Meta-Architecture

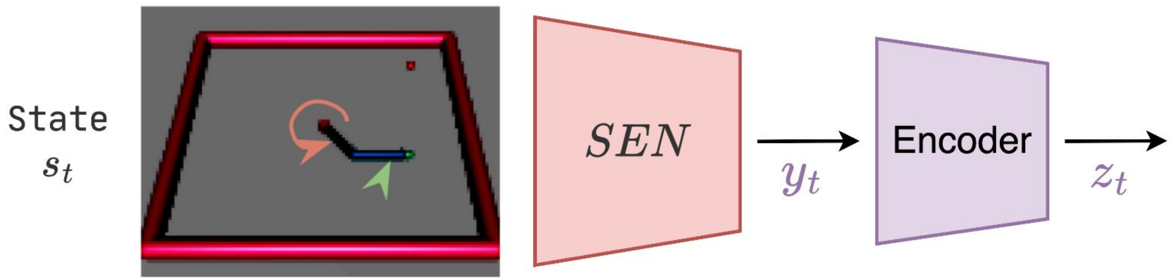
Learn world models with symmetry



feature: known group action, network: equivariant neural network

Meta-Architecture

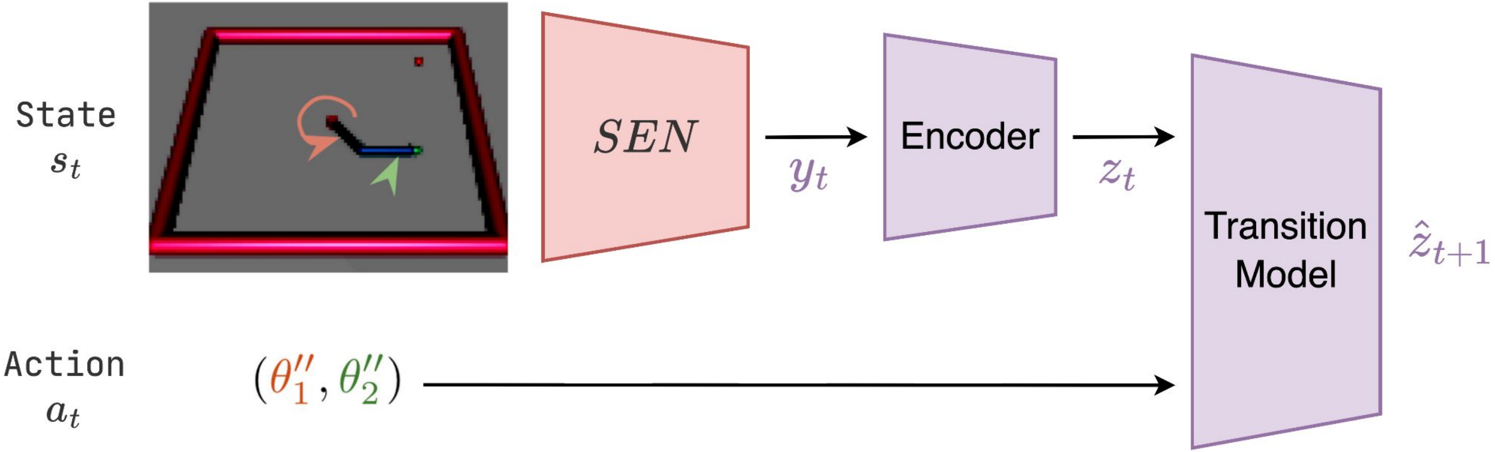
Learn world models with symmetry



feature: known group action, network: equivariant neural network

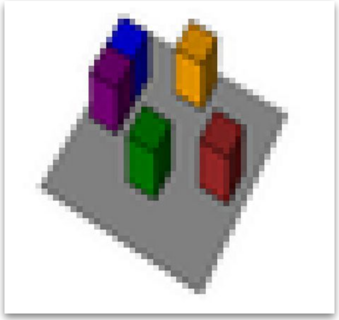
Meta-Architecture

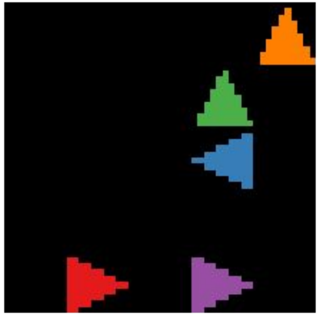
Learn world models with symmetry

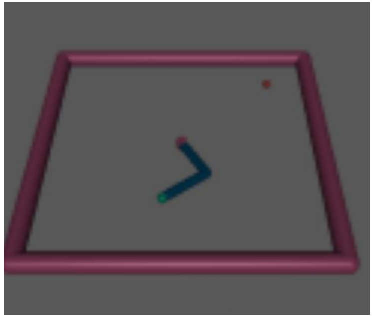


feature: known group action, network: equivariant neural network

Domains







2D Shapes & 3D Blocks	Rush Hour	Reacher	3D Teapot
π/2 rotation Object permutation	π/2 rotation Object permutation	π/2 rotation flip, translation	3D rotation
C_4- equivariant MPNN[1,2]	C_4- equivariant MPNN[1,2]	E(2)-steerable CNN, equivariant MLP[3]	Matrix Multiplication

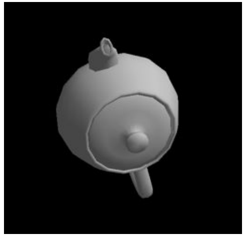
1. Cohen, Taco, and Max Welling. "Group equivariant convolutional networks." *International conference on machine learning* (2016).
2. Scarselli, Franco, et al. "The graph neural network model." *IEEE transactions on neural networks* 20.1 (2008): 61-80.
3. Weiler, Maurice, and Gabriele Cesa. "General e (2)-equivariant steerable cnns." *Advances in Neural Information Processing Systems* 32 (2019).z

Model Performance

3D Teapot

Ground truth

Non-equivariant



Ours (SEN)

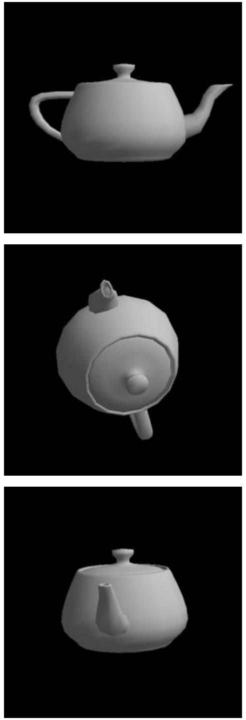
Model Performance

3D Teapot

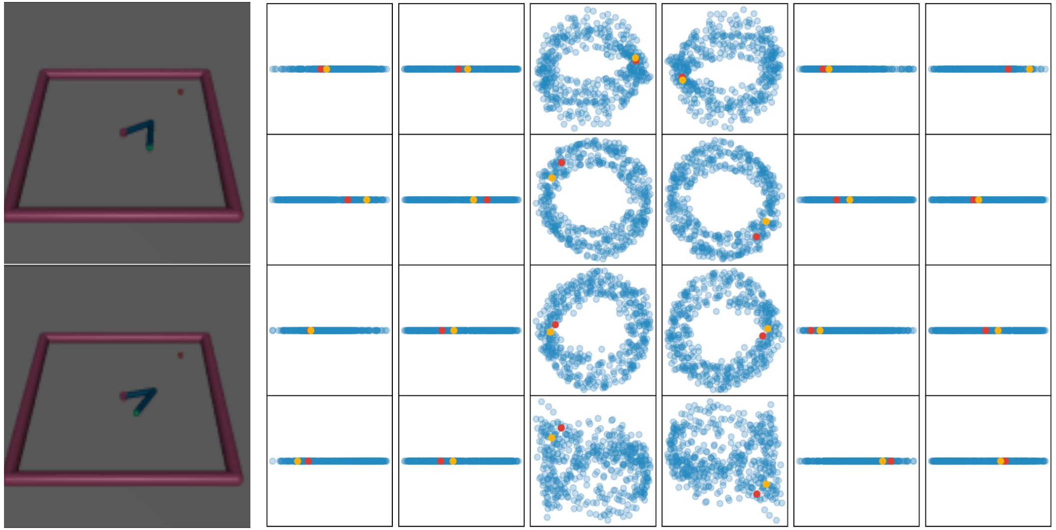
Ground truth

Non-equivariant

Ours (SEN)



Reacher



Generalization over limited actions

Train only on {up,right,down}, test on all {up,right,down,left}

	Limited Actions	Model	MRR (10 step, %)	DIE (10 step)
3D Blocks	{up,right,down}	Non-equivariant	61.8 ± 13.0	181 ± 79.0
		Fully-equivariant	86.0 ± 31.0	15 ± 9.1
		Ours	**100 ± 0.0**	**5 ± 4.7**

SEN achieves higher performance, lower equivariance error than baselines

Generalization benefits of equivariant neural networks

Summary

Group-structured representations

- Learn in latent space
- Connection with disentanglement

Paper 1: Learn autoencoder and parameterize transformation, use entanglement loss

Paper 2: Pair conventional network with equivariant network and train end-to-end