GAUGE EQUIVARIANT CNNS

An Idea From Physics Helps AI See in Higher Dimensions

Gene Olafsen

Presentation Based on:
https://www.quantamagazine.org/an-idea-from-physics-helps-ai-see-in-higher-dimensions-20200109/

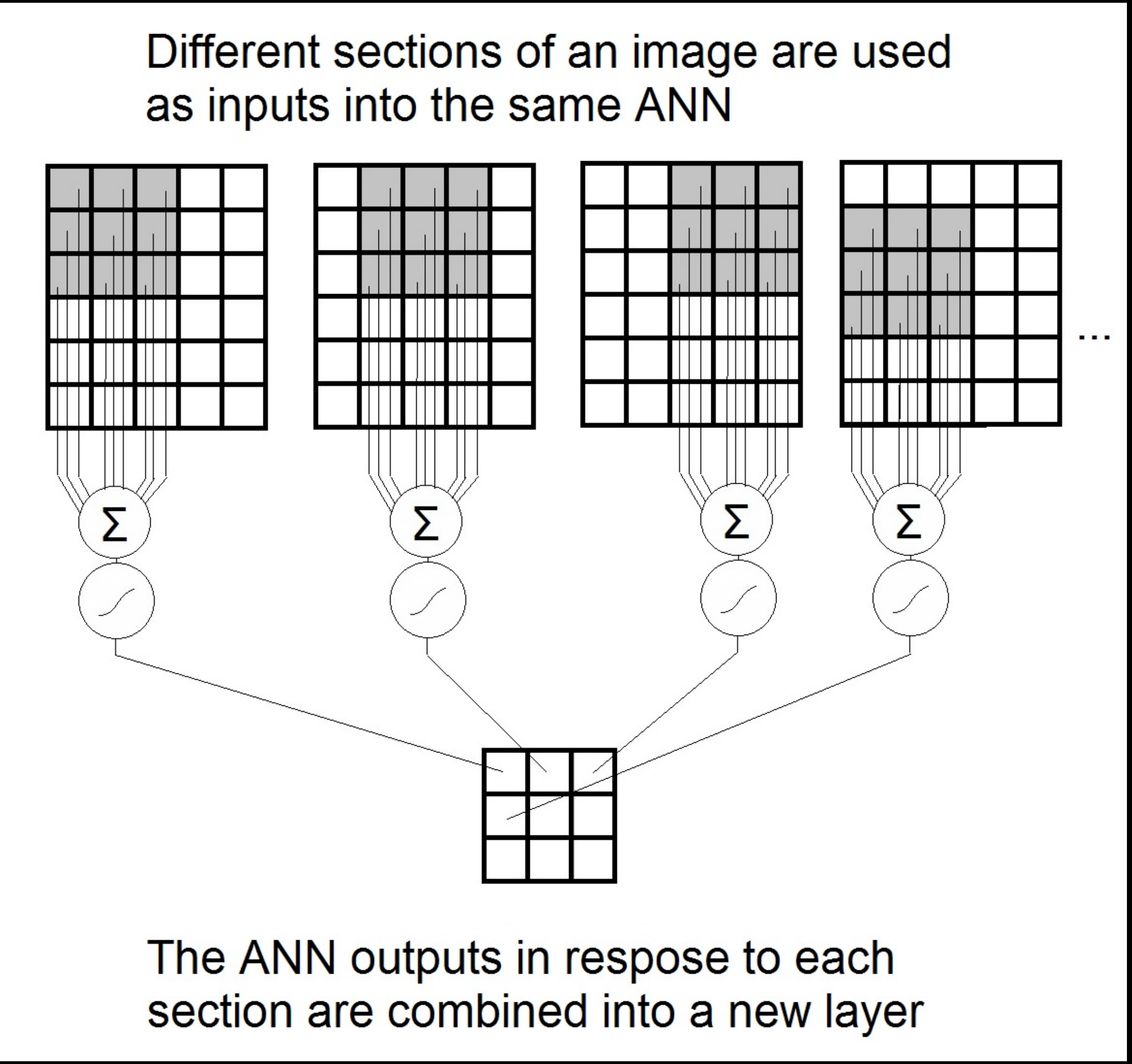
OVERVIEW

- CNNs are good with learning two-dimensional data patterns
  - object identification
  - handwriting

- CNNs are not good
  - point clouds (lidar) evaluation
  - models without a built-in planar geometry

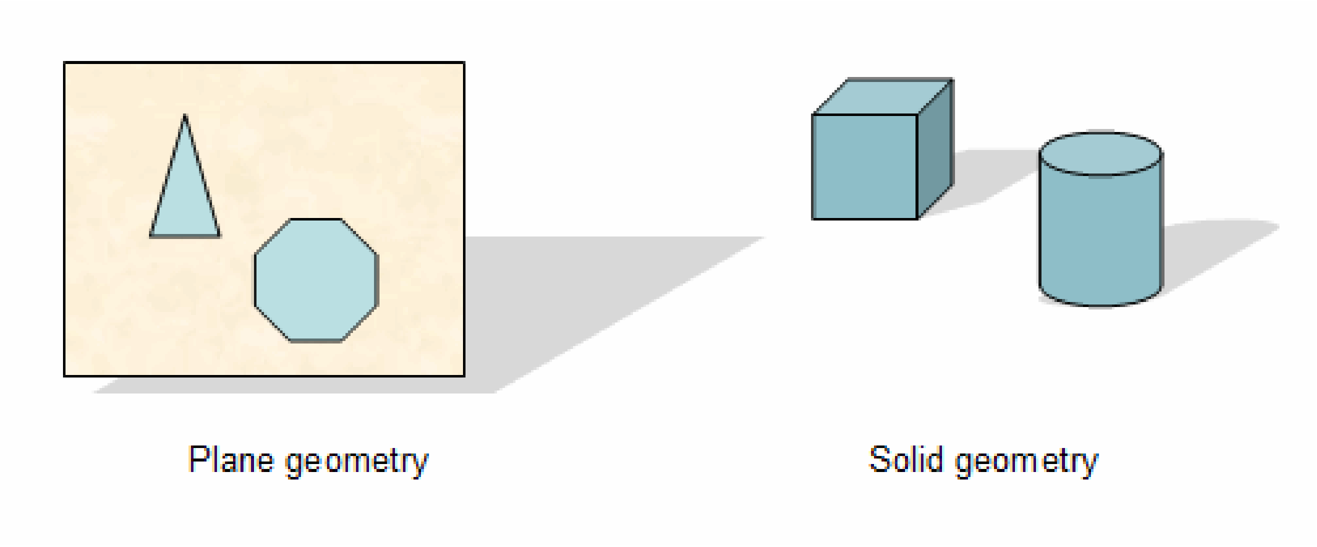
CONVOLUTION

- ANN (Artificial Neural Network)
- A convolution is a linear operation that involves the multiplication of a set of weights with the input.
- The multiplication is performed between an array of input data and a two-dimensional array of weights, called a filter or a kernel.



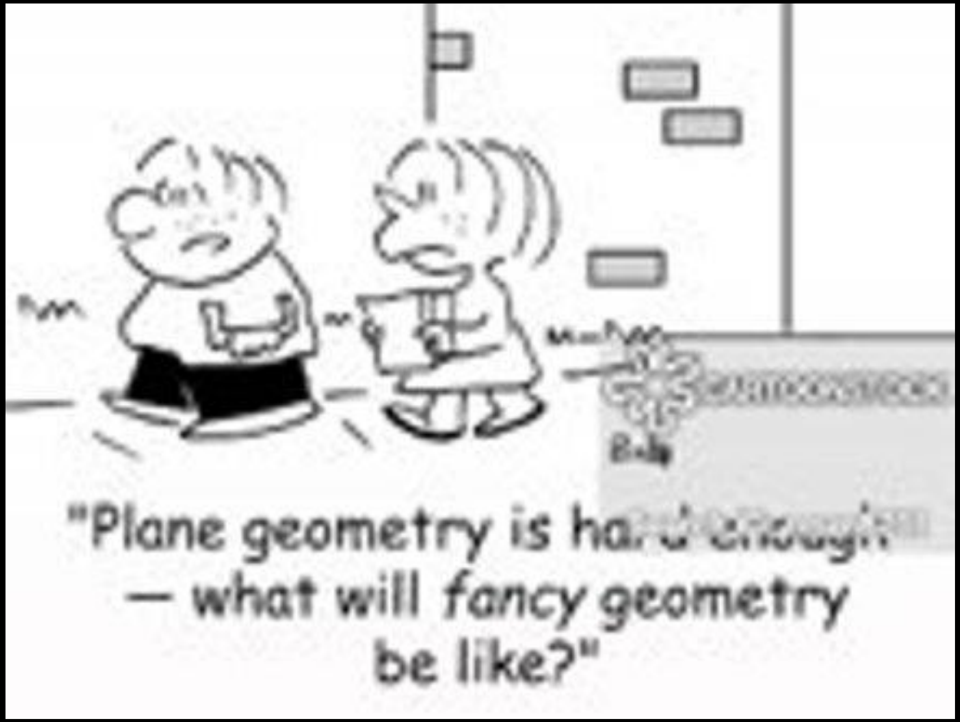
PLANE GEOMETRY

- Geometry can be broken into two broad types:
  - Plane geometry-- deals with only two dimensions
  - Solid geometry-- which allows all three



PLANE GEOMETRY

- In plane geometry, all the shapes exist in a flat plane.
  - A plane can be thought of an a flat sheet with no thickness, and which goes on for ever in both directions.
  - The plane is absolutely flat and infinitely large.



LIFTING CNNS OUT OF "FLATLAND"

- The current buzz around CNNs are by and large describing advancements that are maniuplating/processing data on 2D planes.
- What is needed is a framework for building neural networks that can learn patterns on any kind of geometric surface.

GAUGE CNNS

- Introducing "gauge-equivariant convolutional neural networks" or gauge CNNs
  - Developed at the University of Amsterdam and Qualcomm AI Research
  - By: Taco Cohen, Maurice Weiler, Berkay Kicanaoglu and Max Welling,

THE PHYSICS OF GAUGE

- Gauge theory, which lies in the center of anything in physics that likes to use the words “quantum” and “field” together.

- A class of quantum field theory, a mathematical theory involving both quantum mechanics and Einstein’s special theory of relativity that is commonly used to describe subatomic particles and their associated wave fields.

GAUGE THEORY PROMISE

- The detectection of patterns not only in a 2D arrays of pixels, but also on spheres and asymmetrically curved objects

PERFORMANCE - KNOWN

- Already, gauge CNNs have greatly outperformed their predecessors in learning patterns in simulated global climate data, which is naturally mapped onto a sphere.

PERFORMANCE - EXPECTED

- The algorithms may also prove useful for improving the vision of drones and autonomous vehicles that see objects in 3D, and for detecting patterns in data gathered from the irregularly curved surfaces of hearts, brains or other organs.

GEOMETRIC DEEP LEARNING

- Michael Bronstein, a computer scientist at Imperial College London, coined the term “geometric deep learning” in 2015 to describe nascent efforts to get off flatland and design neural networks that could learn patterns in nonplanar data.

GEOMETRIC DEEP LEARNING - EFFICIENCY

- According to Michael Bronstein, the change also made the neural network dramatically more efficient at learning.
- Standard CNNs “used millions of examples of shapes [and needed] training for weeks.”
- Bronstein used about 100 shapes in different poses and trained for maybe half an hour.

LEARNING WITH LESS

- Separately, Taco Cohen and his colleagues in Amsterdam investigated data efficiency- how to train convolutional neural networks with fewer examples.
- The reduced training set requirement is helpful whether Geometric Deep Learning is utilized or not.
- Specifically, in cancer detection- there are relatively few cancerous nodule images which are labeled, medically accurate and privacy released; in comparison to say 'cat' or 'car' images.

3D ROTO-TRANSLATION GROUP CONVOLUTIONS (G-CONVS)

- In 2016, Cohen and Welling co-authored a paper defining how to encode image variations into a neural network as geometric symmetries-- after all, a lung tumor is still a lung tumor, even if it's rotated or reflected within an image.

- Result: The neural network could identify visual evidence of the disease using just one-tenth of the data used to train other networks.

- This impressive achievement is based on CNNs operating on typical 2D planes.

G-CONVS LIMITATION

- Weiler, Cohen and Welling had extended futher extended their equivariance techniques with CNN's through 2018.
- These approaches still weren't general enough to handle data on manifolds with a bumpy, irregular structures- these are structures that we most frequently encounter in the 'real world'.

GAUGE EQUIVARIANCE EXAMPLE

- Imagine measuring the length of a football field in yards, then measuring it again in meters. The numbers will change, but in a predictable way.
- Similarly, two photographers taking a picture of an object from two different vantage points will produce different images, but those images can be related to each other.
- Gauge equivariance ensures that physicists' models of reality stay consistent, regardless of their perspective or units of measurement.
- Gauge CNNs make the same assumption about data.

EINSTEIN AND EQUIVARIANCE

- Einstein in 1916: “The general laws of nature are to be expressed by equations which hold good for all systems of coordinates.”

CNNS AND MANIFOLDS

- A manifold is a simple thing. Every 2-D surface you see can be considered a manifold. The surface of a sphere, the surface of a cube, all manifolds.
- Easily understood, but now we will get a little more into the weeds and describe a manifold in mathematical terms...
- In mathematics, a manifold is a topological space that locally resembles Euclidean space near each point.

TOPOLOGICAL SPACE

- A topological space may be defined as a set of points, along with a set of neighbourhoods for each point, satisfying a set of axioms relating points and neighbourhoods. The definition of a topological space relies only upon set theory and is the most general notion of a mathematical space that allows for the definition of concepts such as continuity, connectedness, and convergence.

EUCLIDEAN SPACE

- Euclidean n-space, sometimes called Cartesian space or simply n-space, is the space of all n-tuples of real numbers, $(x_1, x_2, ..., x_n)$.

MANIFOLDS...

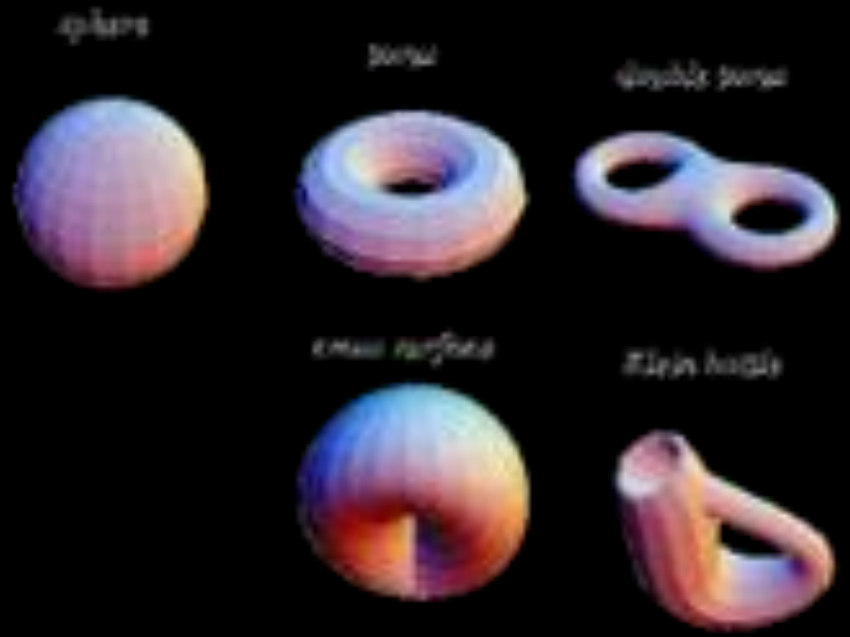
- Let's talk about manifolds:
  - One Dimensional
  - Two Dimensional
  - and beyond...

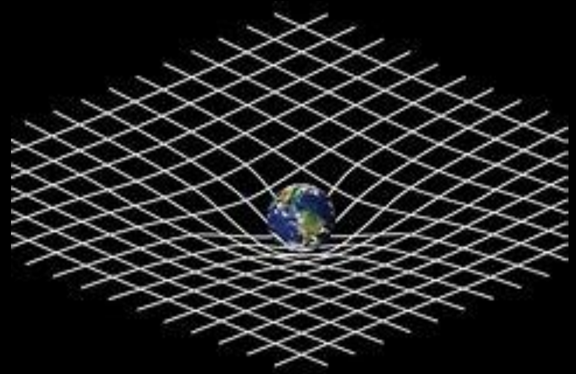
ONE DIMENSIONAL MANIFOLDS

- One-dimensional manifolds include lines and circles.
- Not figure-eights because of the 'crossing point'.
- Circles, parabolas, hyperbolas and cubic curves are all 1D Manifolds.
- Note: the four different colours are all on separate axes and extend out to infinity if it has an open end

TWO DIMENSIONAL MANIFOLDS

- The simplest two dimensional manifold is a sphere.
- It can be imagined that each infinitesimal patch of the sphere locally resembles a 2D Euclidean plane.
- Similarly, any 2D surface (including a plane) that doesn't self-intersect is also a 2D manifold.
- A great analogy illustrating the fact that two dimensional manifolds locally resemble a 2D plane is Earth. We know that the Earth is round but when we stand in a field it looks flat.





BTW

4-D Space-time is a manifold.

THE MANIFOLD ADVANTAGE

- The concept of a manifold is central to many parts of geometry and modern mathematical physics because it allows complicated structures to be described and understood in terms of the simpler local topological properties of Euclidean space.

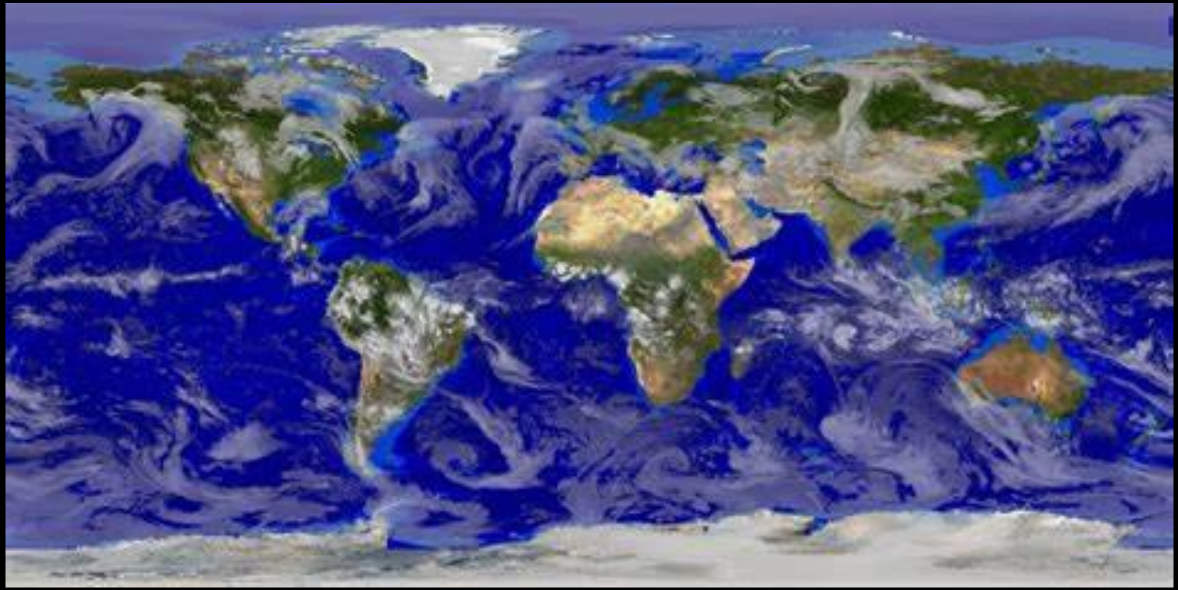
STICKING TOGETHER (PATCHWORK)

- A manifold can be constructed by gluing together pieces in a consistent manner, making them into overlapping charts.
- Surgery theory is a collection of techniques used to produce one finite-dimensional manifold from another in a 'controlled' way, introduced by John Milnor (1961).

EXAMPLE: WORLD WEATHER FORECASTING

- Weather prediction using CNNs for a single country is straightforward. The country has North, South, East and West borders.

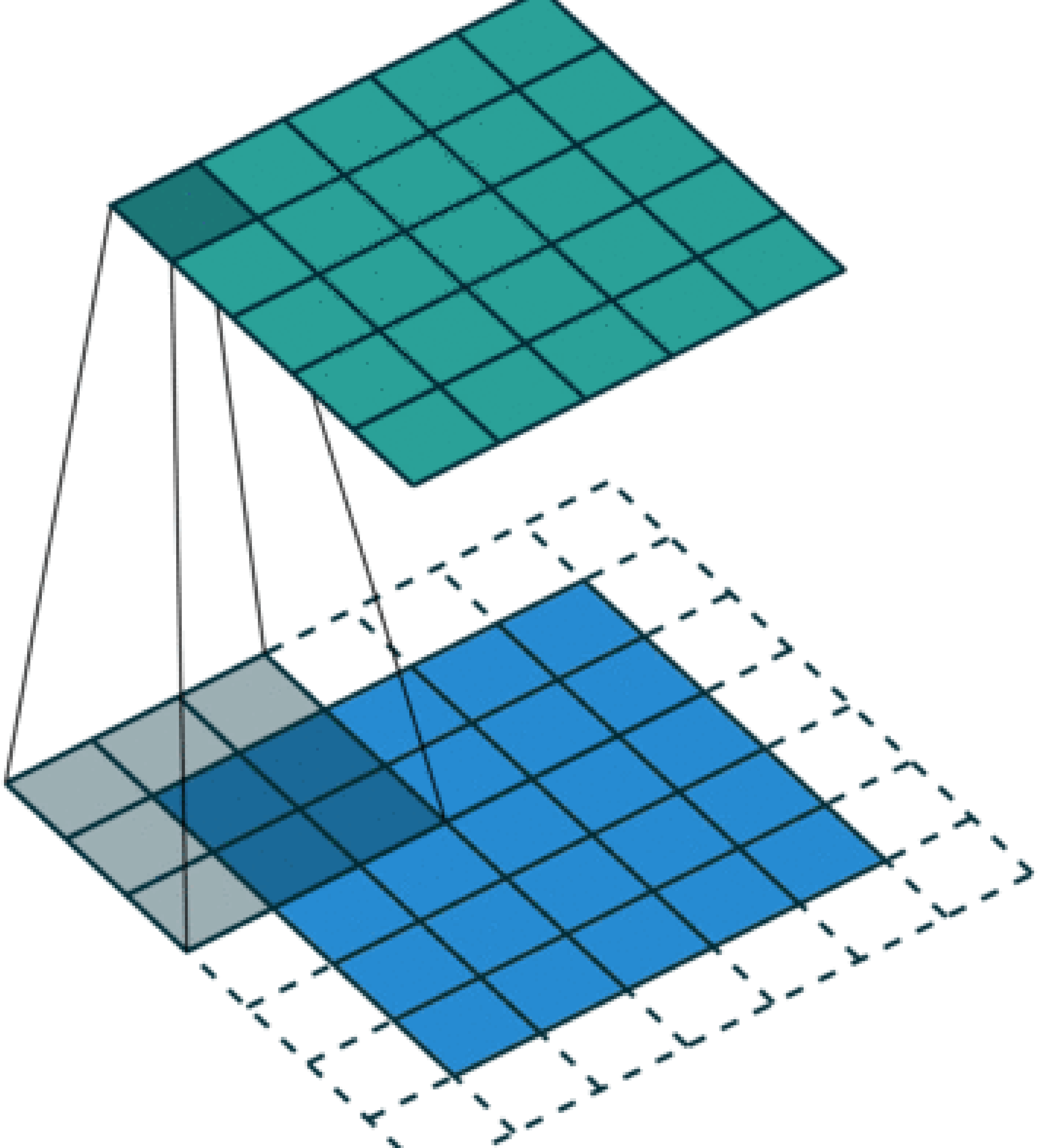
HOW WOULD YOU MODEL WEATHER FOR THE GLOBE?



- There is a problem. The left and the right edge are the same spot in reality. Further, the entire top edge corresponds to a single point, as does the lower edge.
- The whole thing is distorted.

CNN ON A PLANE

- The convolution neural network processing approach only works on a plane.

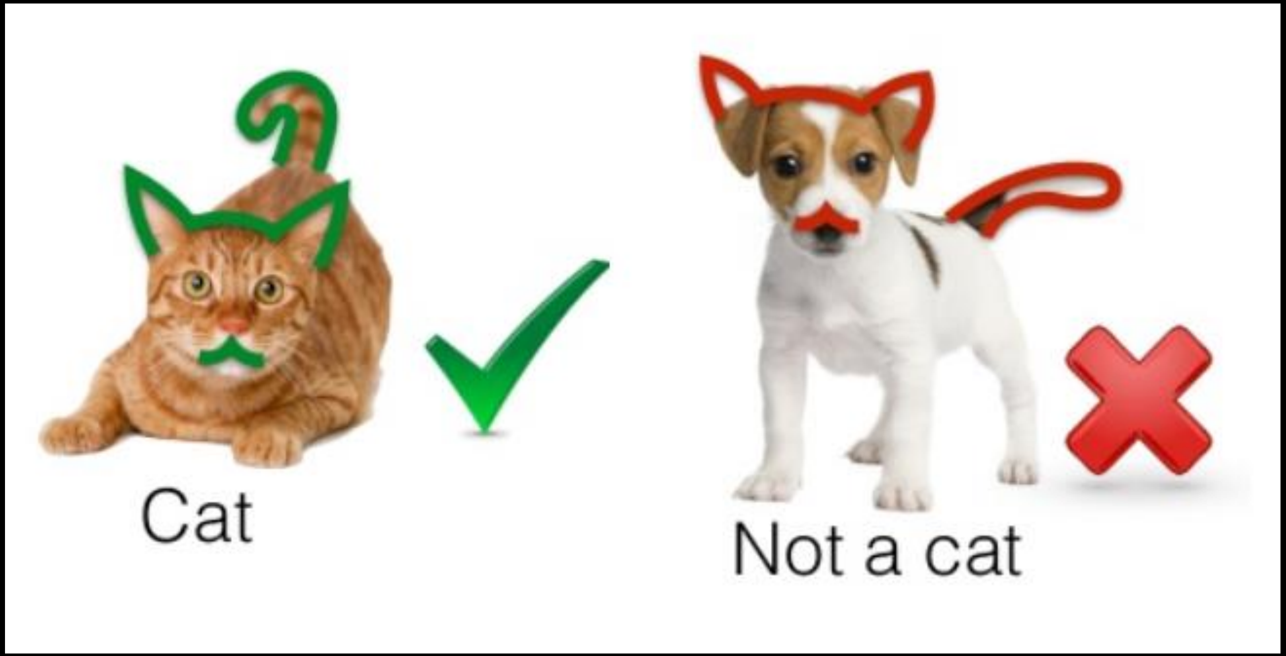


POWER WINDOWS

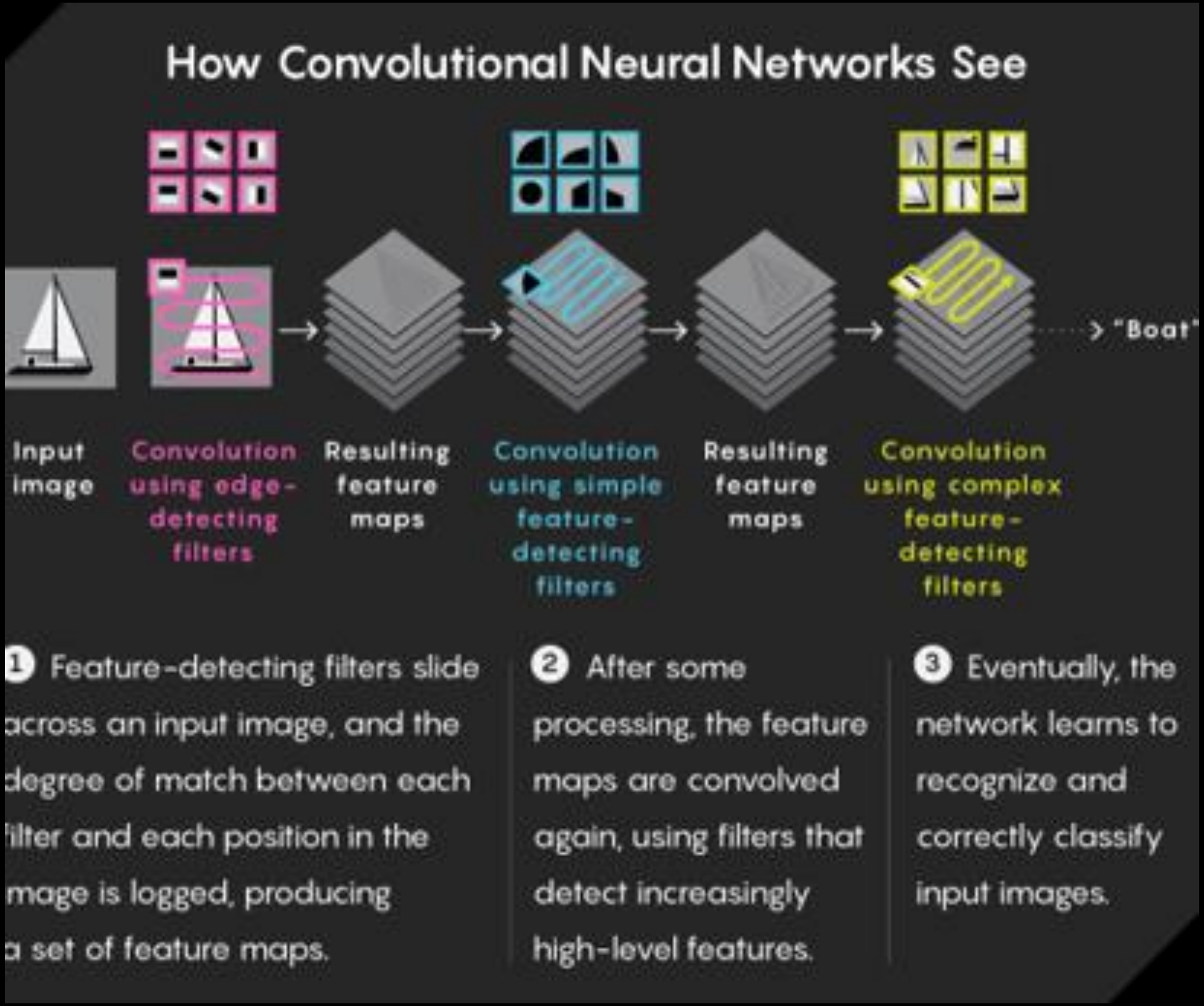
- We have seen that "convolution," is like a sliding window, lets a layer of the neural network perform a mathematical operation on small patches of the input data and then pass the results to the next layer in the network. A convolutional neural network "slides" many of these "windows" over the data like filters to detect patterns.

CAT FEATURES

- If we are working on a CNN to detect cats- lower level layers may detect "edges" whereas higher level layers extract hight-level features, such as eyes, ears or tail. In a fully trained model, the results of the layered convolutions will assign a label of "cat" or "not cat".

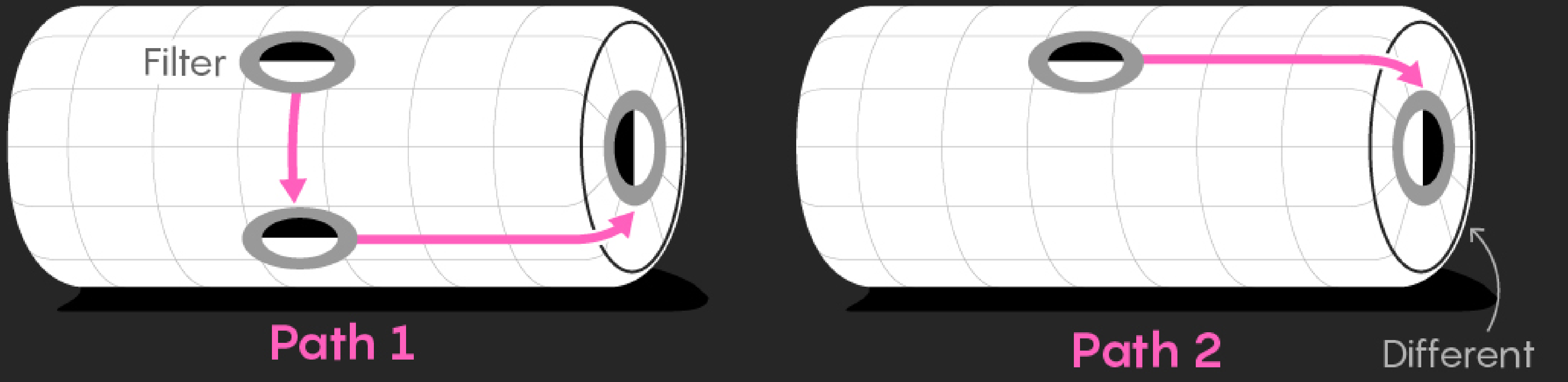


- CNN processing on a curved surface will result in processing images with projection distortions- much like tracing an image onto drafting velum held over the curved surface. (Think of Greenland's distortion on a Mercator projection map.)



How Convolutional Neural Networks See
"Boat"
Input image
Convolution using edge-detecting filters
Resulting feature maps
Convolution using simple feature-detecting filters
Resulting feature maps
Convolution using complex feature-detecting filters
1 Feature-detecting filters slide across an input image, and the degree of match between each filter and each position in the image is logged, producing a set of feature maps.
2 After some processing, the feature maps are convolved again, using filters that detect increasingly high-level features.
3 Eventually, the network learns to recognize and correctly classify input images.

Convolution on Curves



Standard convolutional neural networks don't work on curved surfaces. An edge-detecting filter that slides across a curved surface will end up in different orientations depending on its path, and the filter will yield different feature maps in each case.

FILTER MOVEMENT

- Slide it up, down, left or right on a flat grid, and it will always stay right-side up.
- Slide it to the same spot by moving over the sphere’s north pole, the filter is now upside down — dark blob on the right, light blob on the left.
- Move the filter around a more complicated manifold, and it could end up pointing in any number of inconsistent directions.

FORGET FILTER ORIENTATION

- The key, explained Welling, is to forget about keeping track of how the filter's orientation changes as it moves along different paths.
- Instead, you can choose just one filter orientation (or gauge), and then define a consistent way of converting every other orientation into it.

FREE LUNCH

- Cohen, Weiler and Welling encoded gauge equivariance — the ultimate "free lunch" — into their convolutional neural network in 2019.
- They did this by placing mathematical constraints on what the neural network could "see" in the data via its convolutions; only gauge-equivariant patterns were passed up through the network's layers.
- "Basically you can give it any surface" — from Euclidean planes to arbitrarily curved objects, including exotic manifolds like Klein bottles or four-dimensional space-time — "and it's good for doing deep learning on that surface," said Welling.

A WORKING THEORY

- The theory of gauge-equivariant CNNs is so generalized that it automatically incorporates the built-in assumptions of previous geometric deep learning approaches — like rotational equivariance and shifting filters on spheres.
- Even Michael Bronstein’s earlier method, which let neural networks recognize a single 3D shape bent into different poses, fits within it. “Gauge equivariance is a very broad framework. It contains what we did in 2015 as particular settings,” Bronstein said.

WEATHER

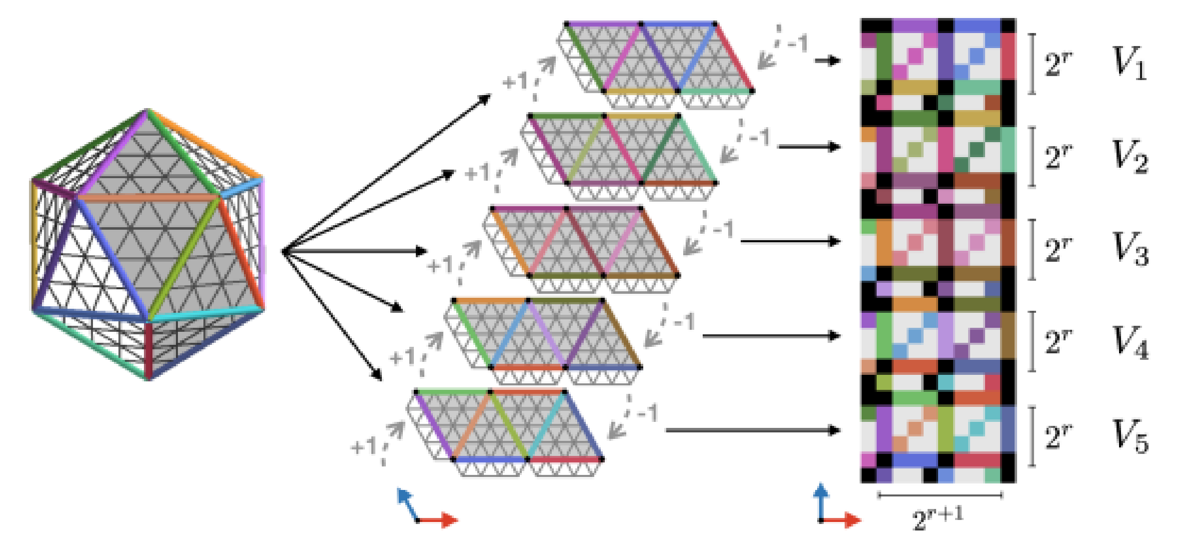
- They used their gauge-equivariant framework to construct a CNN trained to detect extreme weather patterns, such as tropical cyclones, from climate simulation data.
- In 2017, government and academic researchers used a standard convolutional network to detect cyclones in the data with 74% accuracy.
- In 2019, the gauge CNN detected the cyclones with 97.9% accuracy. (It also outperformed a less general geometric deep learning approach designed in 2018 specifically for spheres — that system was 94% accurate.)

ICOSAHEDRON

- Example of gauge equivariant convolution on the icosahedron.
- The very special shape of this manifold makes it possible to implement gauge equivariant convolution in a way that is both numerically convenient (no interpolation is required), and computationally efficient (the heavy lifting is done by a single conv2d call).
- The icosahedron is a regular solid with 20 faces, 30 edges, and 12 vertices. It has 60 rotational symmetries.

ICOSAHEDRON GAUGE

Figure 4. The Icosahedron with grid $\mathcal{H}_r$ for $r = 2$ (left). We define 5 overlapping charts that cover the grid (center). Chart V_5 is highlighted in gray (left). Colored edges that appear in multiple charts are to be identified. In each chart, we define the gauge by the standard axis aligned basis vectors $e_1, e_2 \in V_i$. For points $p \in U_i \cap U_j$, the transition between charts involves a change of gauge, shown as $+1 \cdot 2\pi/6$ and $-1 \cdot 2\pi/6$ (elements of $G = C_6$). On the right we show how the signal is represented in a padded array of shape $5 \cdot (2^r + 2) \times (2^{r+1} + 2)$.



MNIST ON AN ICOSAHEDRON

- In order to validate the implementation, highlight the potential benefits of the method, and determine the necessity of each part of the algorithm, the researchers perform a number of experiments with the MNIST dataset, projected to the icosahedron.

MNIST DATA PREPARATION

- The researchers generated three different versions of the training and test sets, differing in the transformations applied to the data.
- In the N condition, No rotations are applied to the data.
- In the I condition, they apply all 60 Icosahedral symmetries (rotations) to each digit.
- Finally, in the R condition, they apply 60 random continuous rotations g ∈ SO(3) to each digit before projecting. A

CNN MODEL

- The main model consists of 7 convolution layers and 3 linear layers.
- The first layer is a scalar-to-regular gauge equivariant convolution layer, and the following 6 layers are regular-to-regular layers.
- These layers have 8, 16, 16, 24, 24, 32, 64 output channels, and stride 1, 2, 1, 2, 1, 2, 1, respectively

RESULTS

- Icosahedral CNNs achieve excellent performance with a test accuracy of up to 99.43%, which is a strong result even on planar MNIST, for non-augmented and non-ensembled models.

Arch.	N/N	N/I	N/R	I/ I	I / R	R / R
S2CNN	99.38	99.38	99.38	99.12	99.13	99.12
NP+NE	99.29	25.50	16.20	98.52	47.77	94.19
NE	99.42	25.41	17.85	98.67	60.74	96.83
NP	99.27	36.76	21.4	98.99	61.62	97.87
S2S	97.81	97.81	55.64	97.72	58.37	89.92
S2R	98.99	98.99	59.76	98.62	55.57	98.74
R2R	**99.43**	**99.43**	**69.99**	**99.38**	**66.26**	**99.31**

Table 1. IcoMNIST test accuracy (%) for different architectures and train / test conditions (averaged over 3 runs). See text for explanation of labels.

BEYOND

- Risi Kondor, a former physicist who now studies equivariant neural networks, said the potential scientific applications of gauge CNNs may be more important than their uses in AI.

REFENCES

- https://www.quantamagazine.org/an-idea-from-physics-helps-ai-see-in-higher-dimensions-20200109/
- https://arxiv.org/abs/1804.04656
- https://arxiv.org/pdf/1902.04615.pdf
- https://towardsdatascience.com/an-easy-guide-to-gauge-equivariant-convolutional-networks-9366fb600b70
- https://en.wikipedia.org/wiki/Manifold
- http://bjlkeng.github.io/posts/manifolds/
- https://www.mathopenref.com/planegeometry.html
- http://mathworld.wolfram.com/EuclideanSpace.html