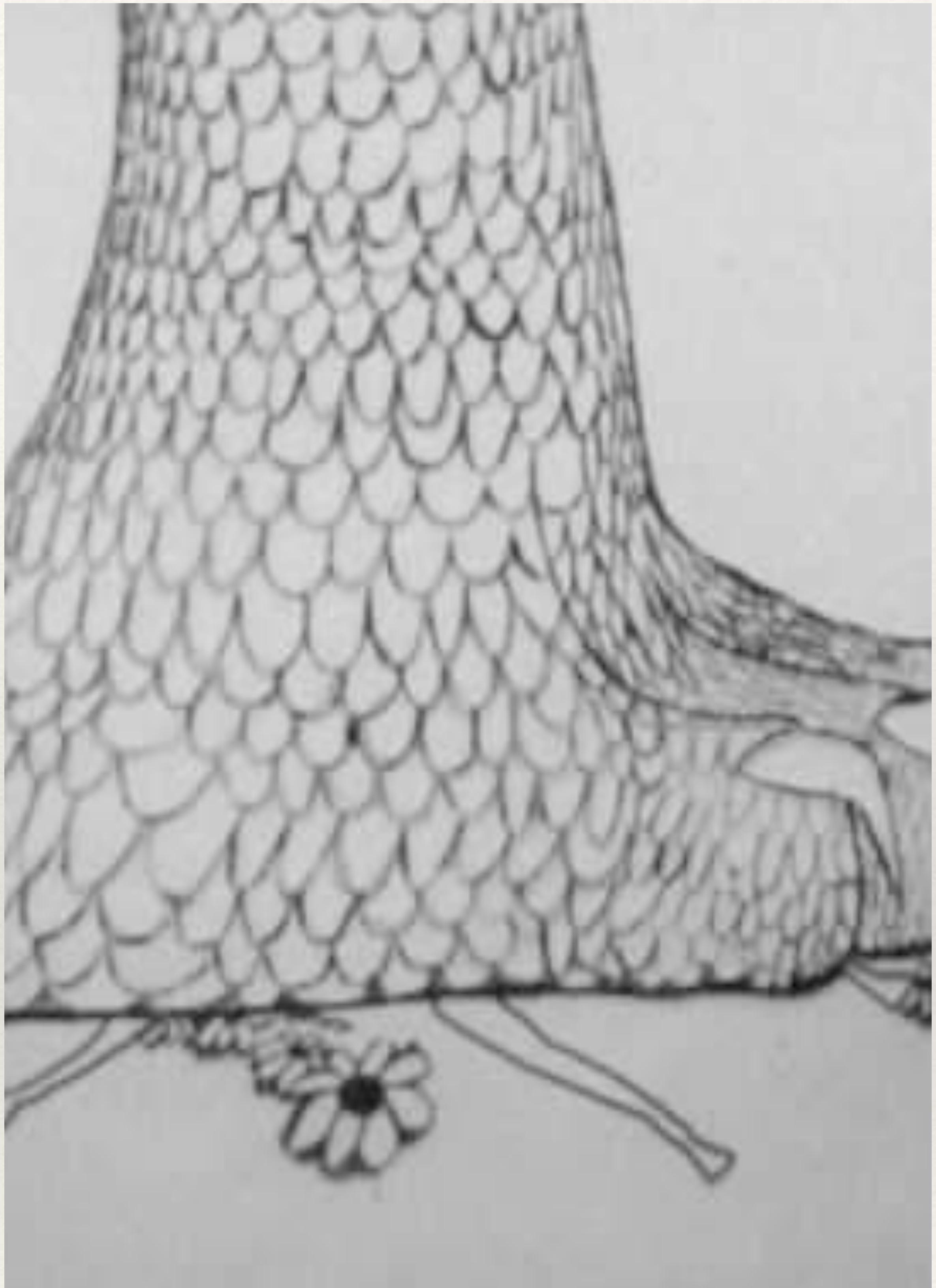
Symmetry meets AI

Veronica Sanz
Universitat de Valencia - IFIC (Spain)
& Sussex University (UK)
with Gabriela Barenboim and Johannes Hirn

arXiv:2103.06115 [cs.LG]



Bambi Meets Godzilla (1969)

Today, we will talk about

Human vs Machine Learning
Human surrender
Looking under the hood
From Physics to Art

My aim is
if you know about ML, make you think a bit differently
if you don't, motivate you to have a closer look

Human vs Machine Learning

Human learning

Previous experience

repeat and improve on a task

predict the evolution of a situation

discover unknown relations

choose the option that maximises return

imagine new possibilities

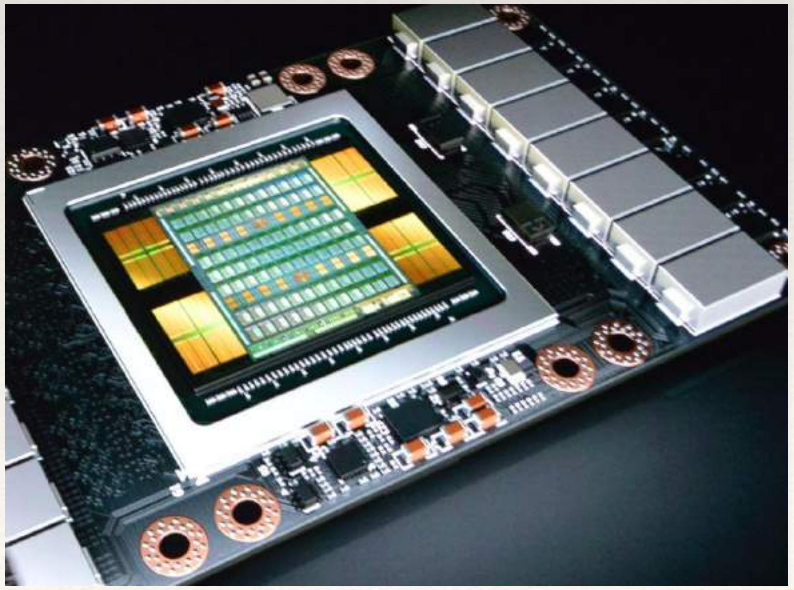
VERY IMPRESSIVE, YET

human learning is limited by
our personal viewpoint,
our collective intelligence (*newspeak?)*
& our inherent capacity to process information
(amount , speed, level of detail)

ON THE OTHER HAND

the ultimate limitations of machine learning
are unknown (if they do exist)
CPU-> GPU, TPU, FPGA, IPU -> …
Quantum Computing, Neurophotonics…

VS



Machine learning

Previous experience

repeat and improve on a task

SUPERVISED MACHINE LEARNING

predict the evolution of a situation

TIME-SERIES LEARNING

discover unknown relations

CLUSTERING/UNSUPERVISED

choose the option that maximises return

REINFORCEMENT LEARNING

imagine new possibilities

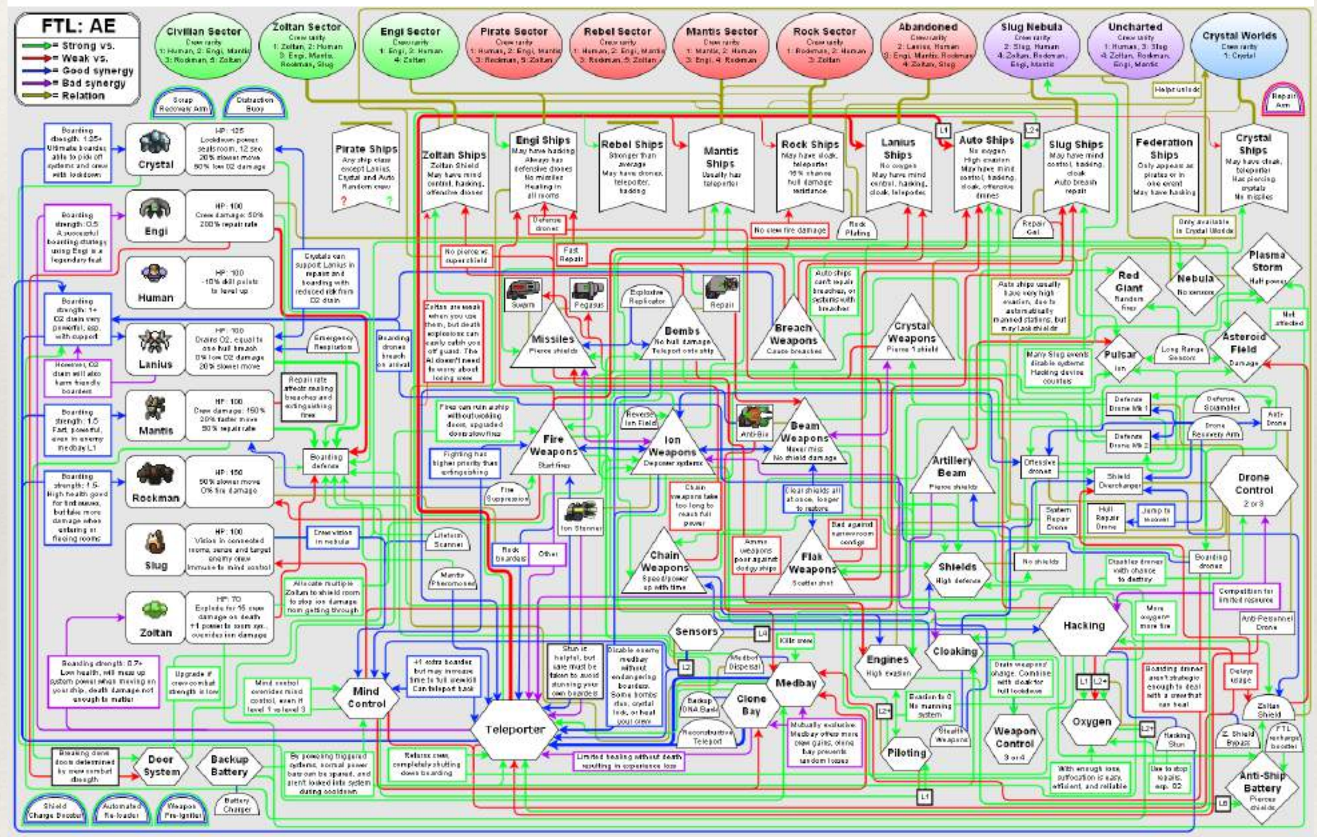
GENERATIVE AI

Nowadays, Machine Learning is in the middle of a revolution: processing speed and storing capacity have increased enormously but **more importantly** the *way* machines learn has changed

TRADITIONALLY

learning was limited to lines of code we (humans) were writing

```
if something_is_in_the_way is True:
    stop_moving()
else:
    continue_moving()
```



we can write *extremely complex* codes and the machine can improve in performing tasks

but the structure of *thought* behind decision making is human

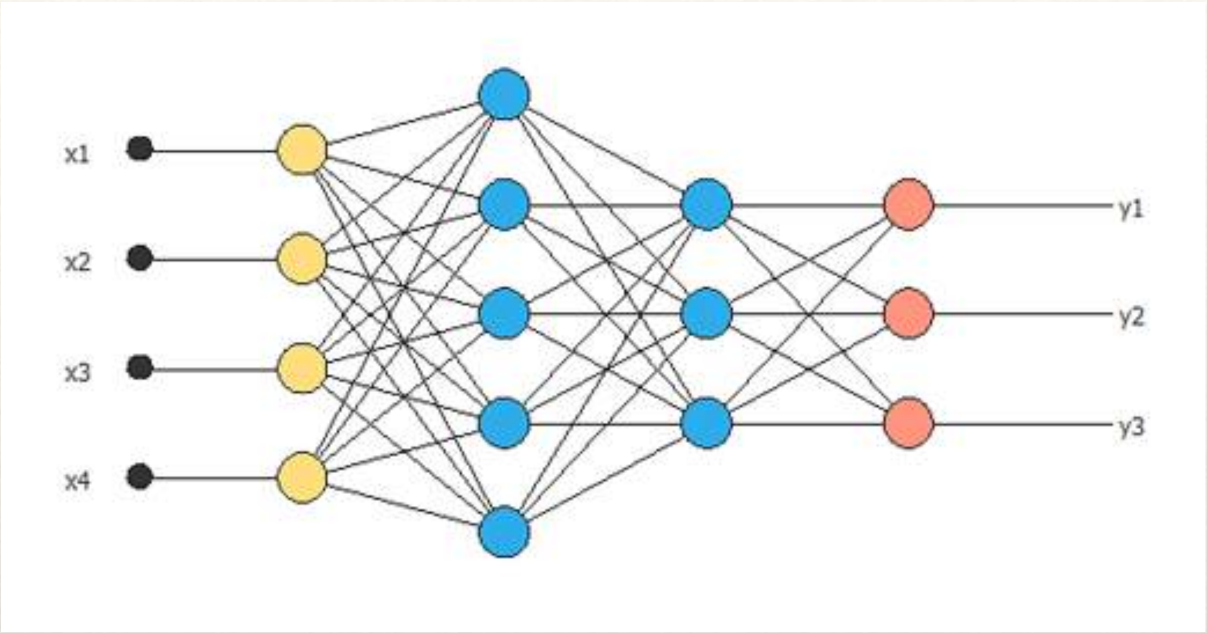
The Machine can't describe relations we haven't coded in

like a born-blind person who is asked to think of ***blue***

A new way of *thinking*: Neural Networks

Structures made of units called *neurons*
and organised by *layers*

inputs



outputs

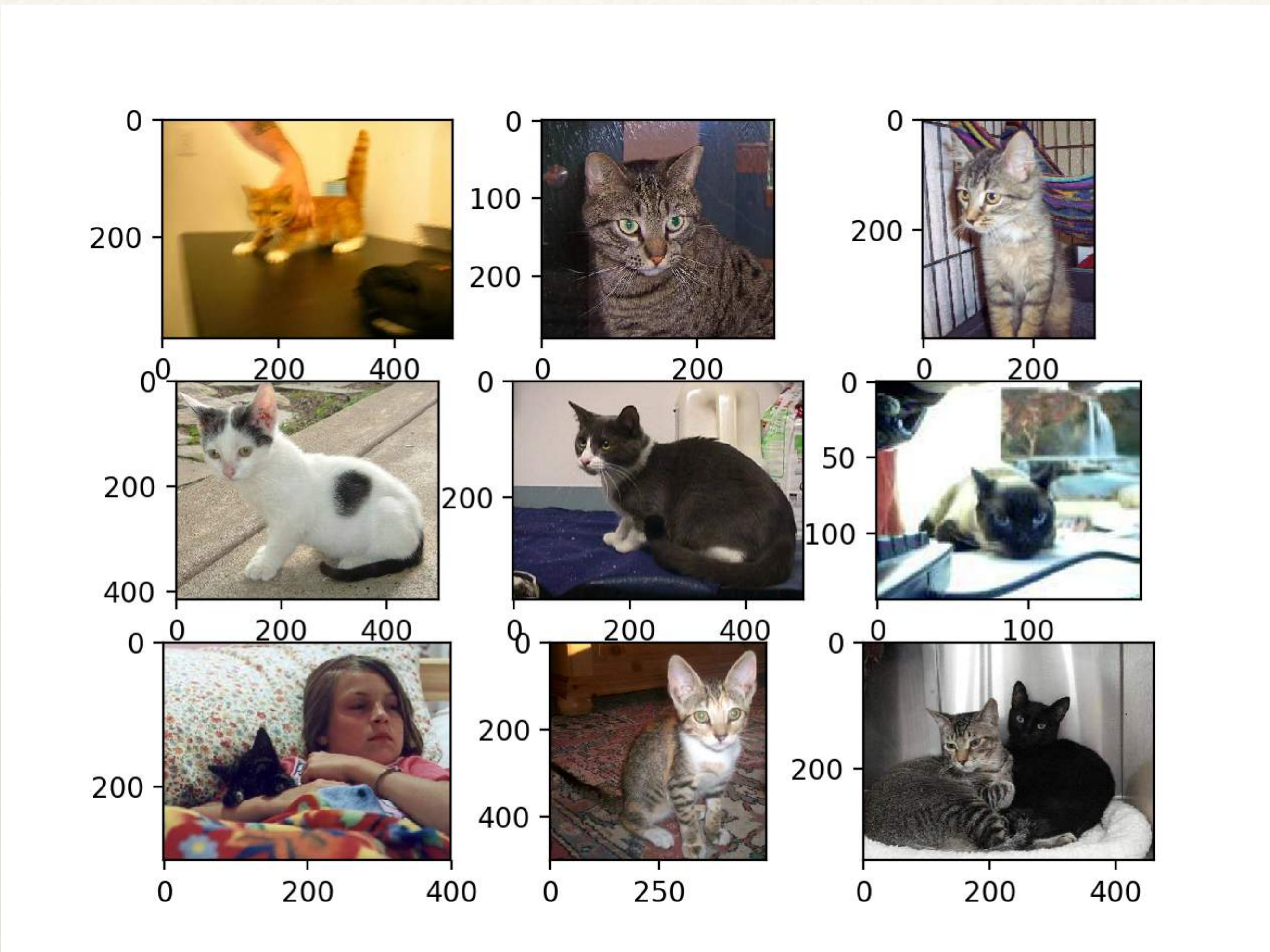
The network learns from data with no structured instructions

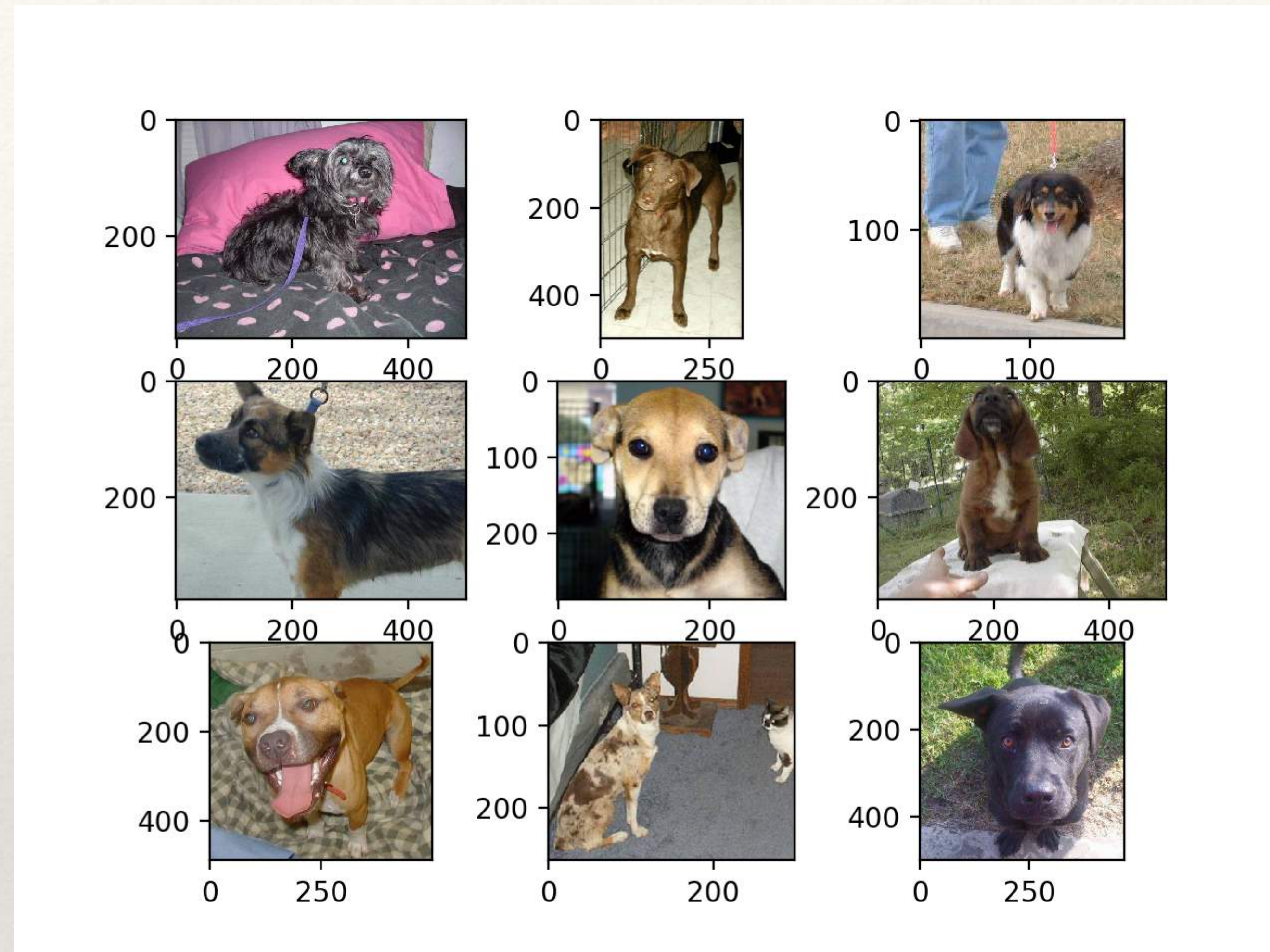
Neural networks are able to explore relations between inputs and outputs which cannot be contained in lines of codes
their degree of expressivity is immense
and it is extremely fast
built from simple units and in a layered architecture

Machines can now tackle really complex tasks

images, speech : are complex

For example: cats/dogs





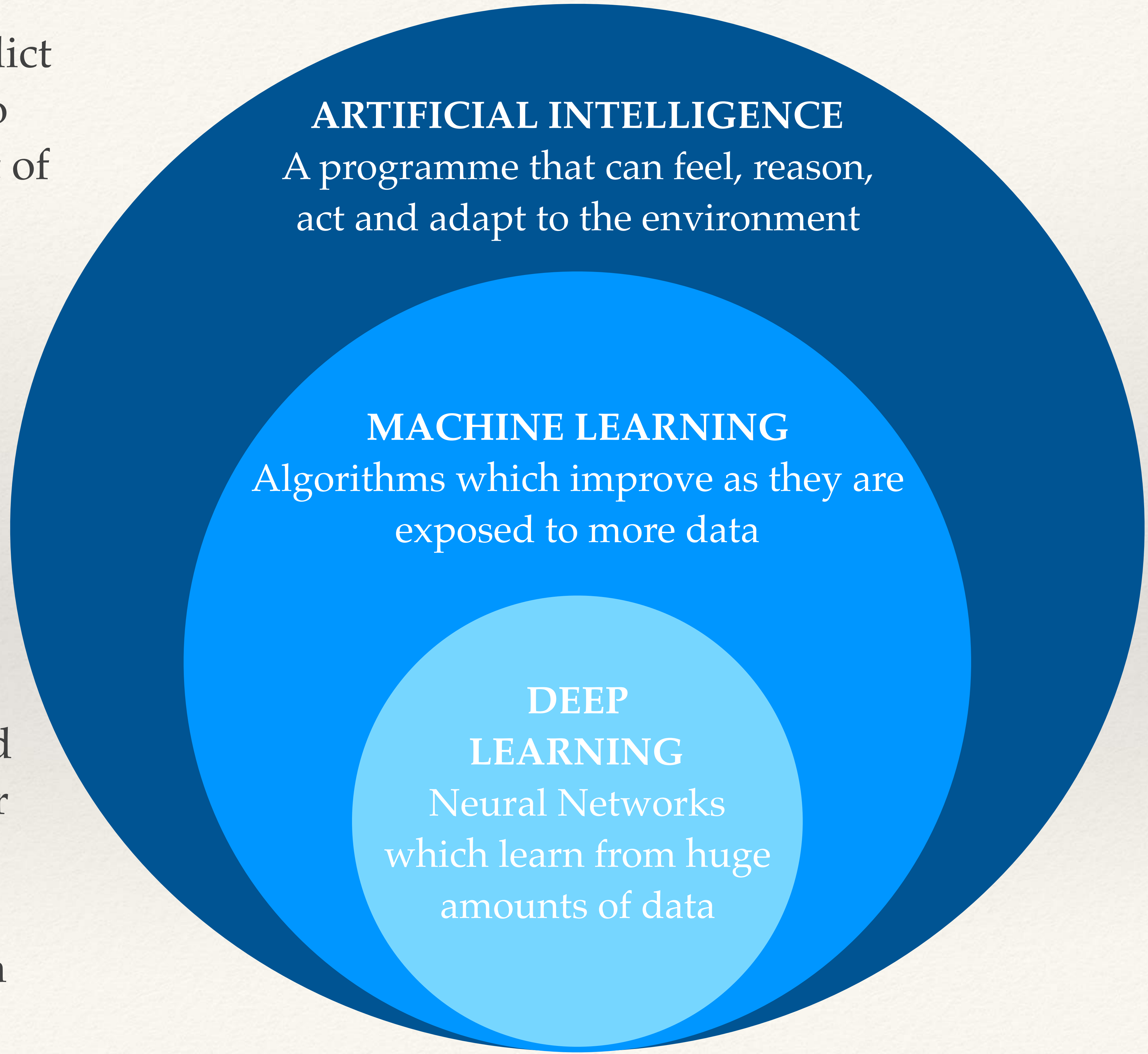
you can distinguish these cats and dogs, right? but how?

would you be able to write a code which classifies them with ~ 100% accuracy? well, a NN can learn to do this!

and many other things, like beating a Go master

This technology is truly *disruptive*

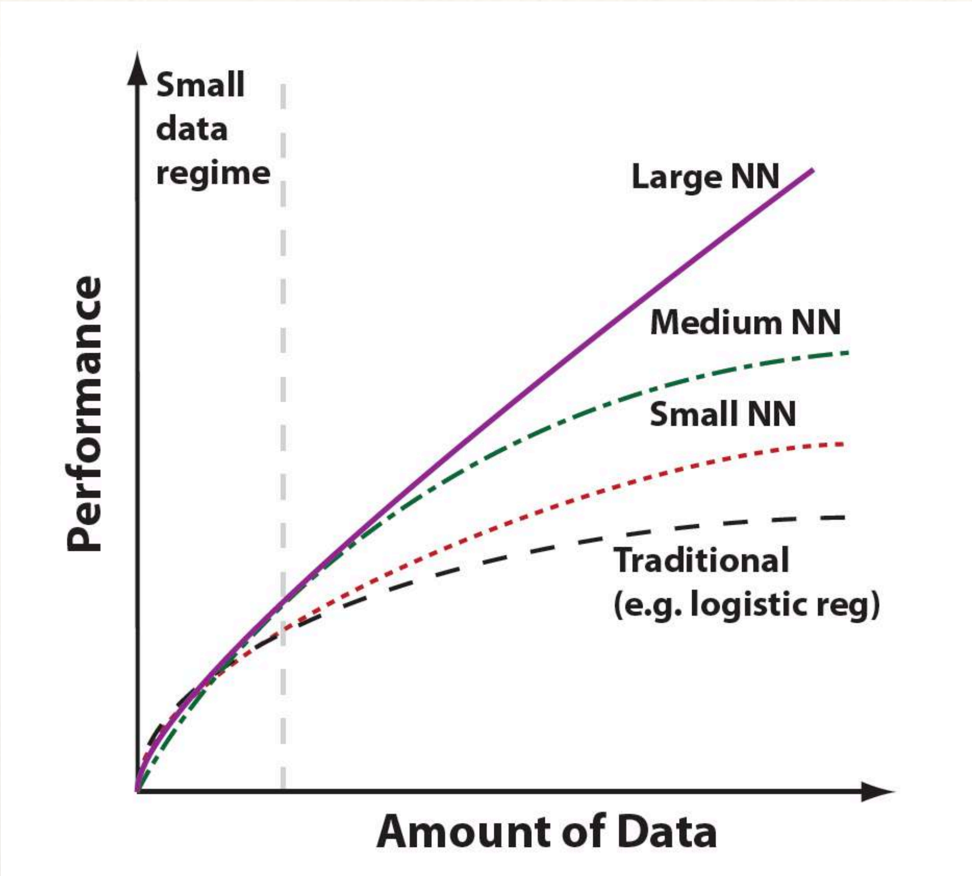
we are unable to predict how fast is going to evolve and the extent of its applications



new algorithms and applications appear every day, and this tendency does not seem to slow down

Human surrender?

Why are NNs so good at learning?



High-bias low-variance, 1803.08823

Good at handling large amounts of data: needle in a haystack

The NN structure (layers, 0/1 gates) allows a high representation power with moderate computational demands, e.g. allows parallelisation, use of GPUs…

It scales better than other learning methods (like SVMs)

Good at learning: ability to learn with little *domain knowledge*

That's something physicists (as humans) are good at

(Physics -> other things)

DNNs are good at this too, they are able to take large streams of data and learn features with little guidance, work like *black boxes*

What's wrong with blackboxes?

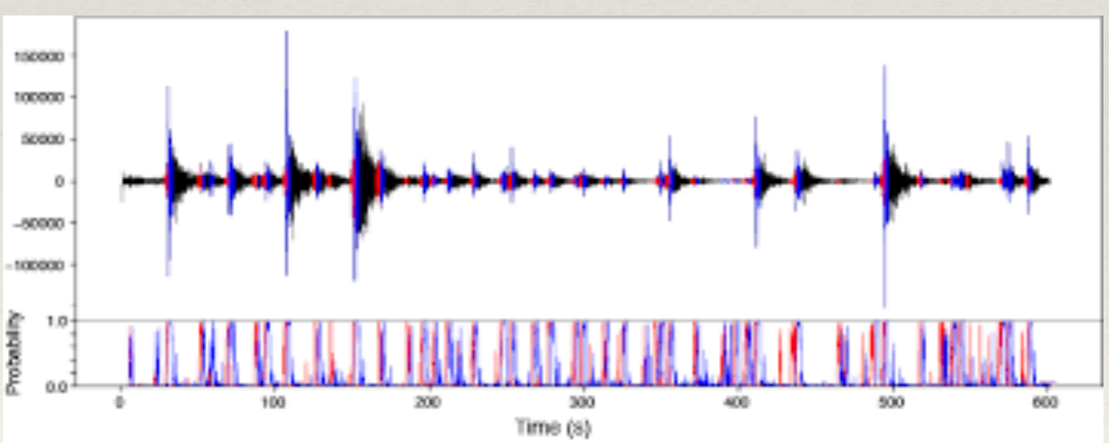
Only open if a disaster happened

If it works, why fix it?
DNN is very powerful, in a way that can be quantified and tensioned against human performance or other techniques

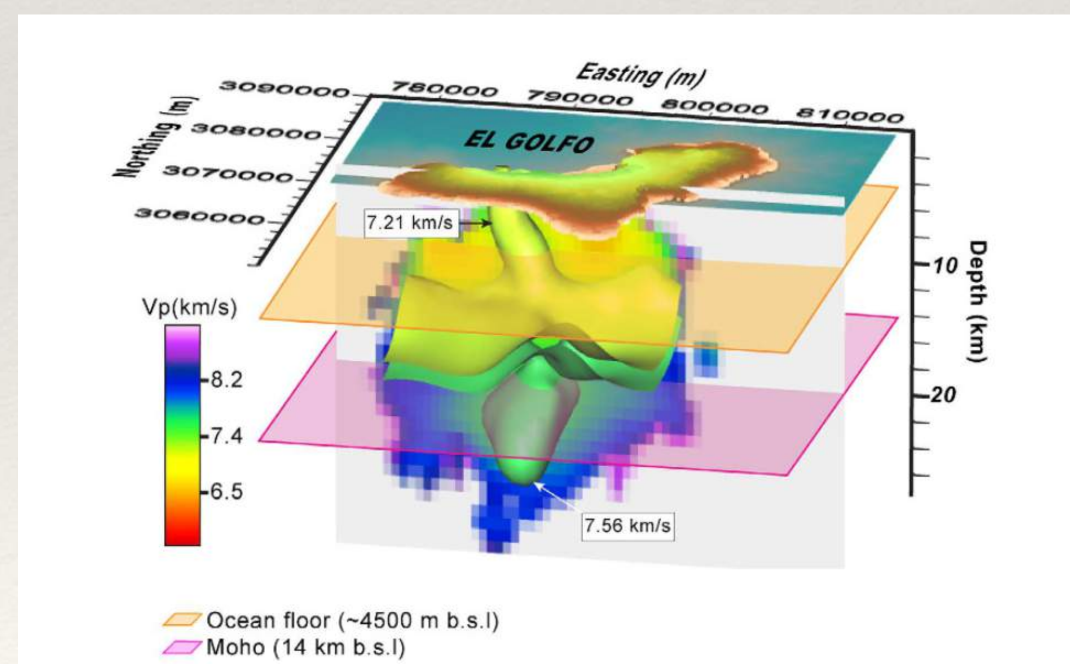
---

Example: collaboration with Seismicity experts

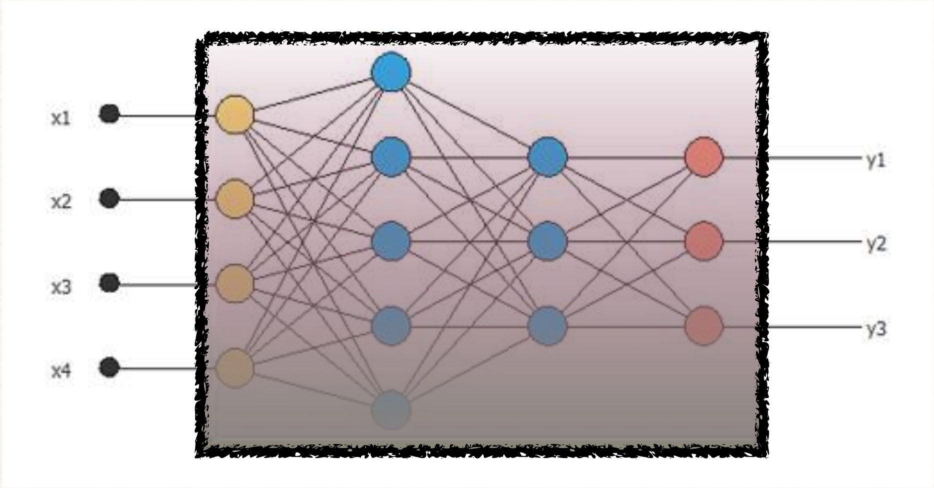
Automatic detection of Earthquakes and phase picking



Tomography



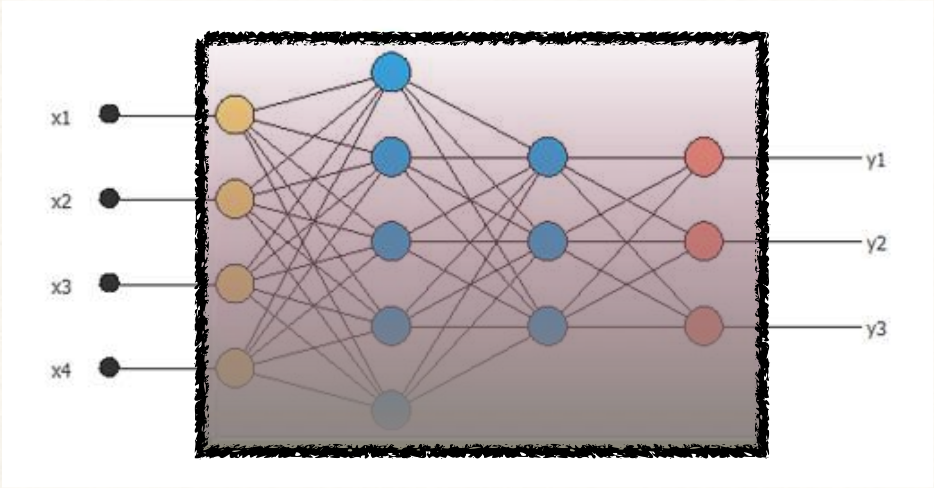
Institut de Ciències del Mar + Universitat de València



What's wrong with blackboxes?

If they do work, and help solve problems?

The lack of understanding hurts our pride as scientists
our job is to understand as much as we humanly can
"If you think you understand quantum mechanics, you don't understand quantum mechanics" R. Feynman, *The Character of Physical Law*

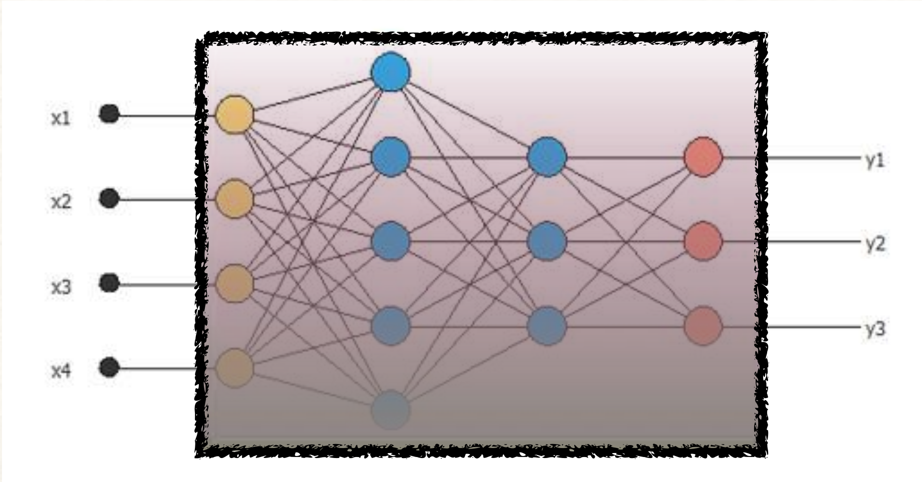


What's wrong with blackboxes?

If they do work, and help solve problems?

The lack of understanding hurts our pride as scientists
our job is to understand as much as we humanly can
"If you think you understand quantum mechanics, you don't understand quantum mechanics" R. Feynman, *The Character of Physical Law*

Any efforts we do to express the workings of NNs from different viewpoints may lead to *new ideas for machine learning*



What's wrong with blackboxes?

If they do work, and help solve problems?

The lack of understanding hurts our pride as scientists
our job is to understand as much as we humanly can
"If you think you understand quantum mechanics, you don't understand quantum mechanics" R. Feynman, *The Character of Physical Law*

Any efforts we do to express the workings of NNs from different viewpoints may lead to *new ideas for machine learning*

The depth and reach of AI in *decision making* is growing very fast
we should be concerned about our lack of control over this
e.g. see EU's draft on regulating AI, April 21st
XAI, Ethical AI… all these require a better understanding of DNNs

Looking under the hood

with symmetries

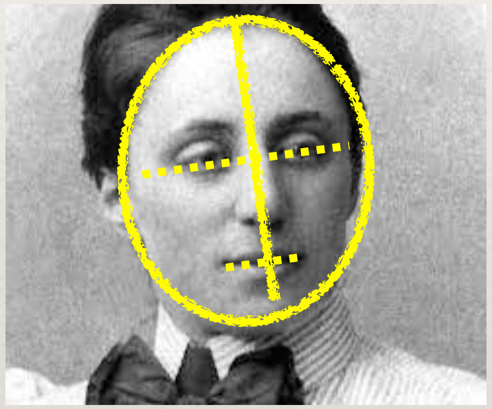
Symmetry is a key concept in Physics

not just as a simplification method or connection with other problems

deeper level: Laws of Physics, understanding of forces, stability…

Symmetries can help with **Machine Learning** problems

e.g., CNNs and data augmentation



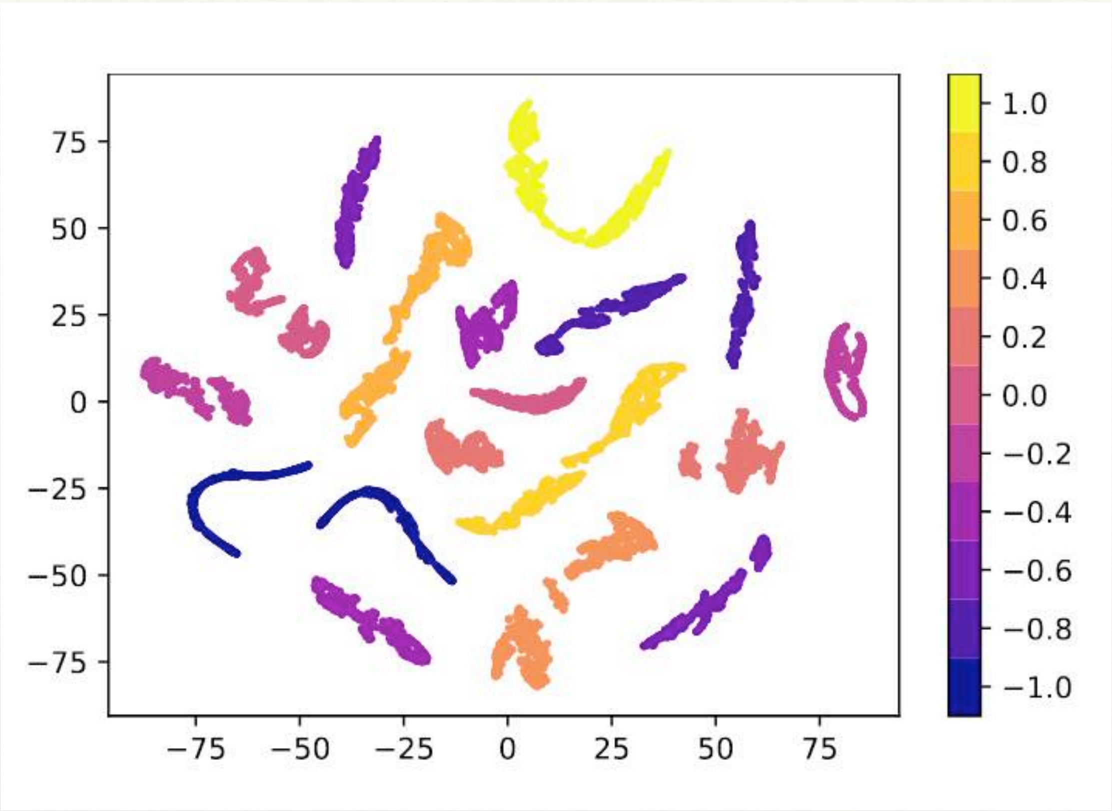
The **concept** of symmetry is part of our shared **human appreciation**

We asked ourselves:

Can Machines Deep-Learn symmetries?

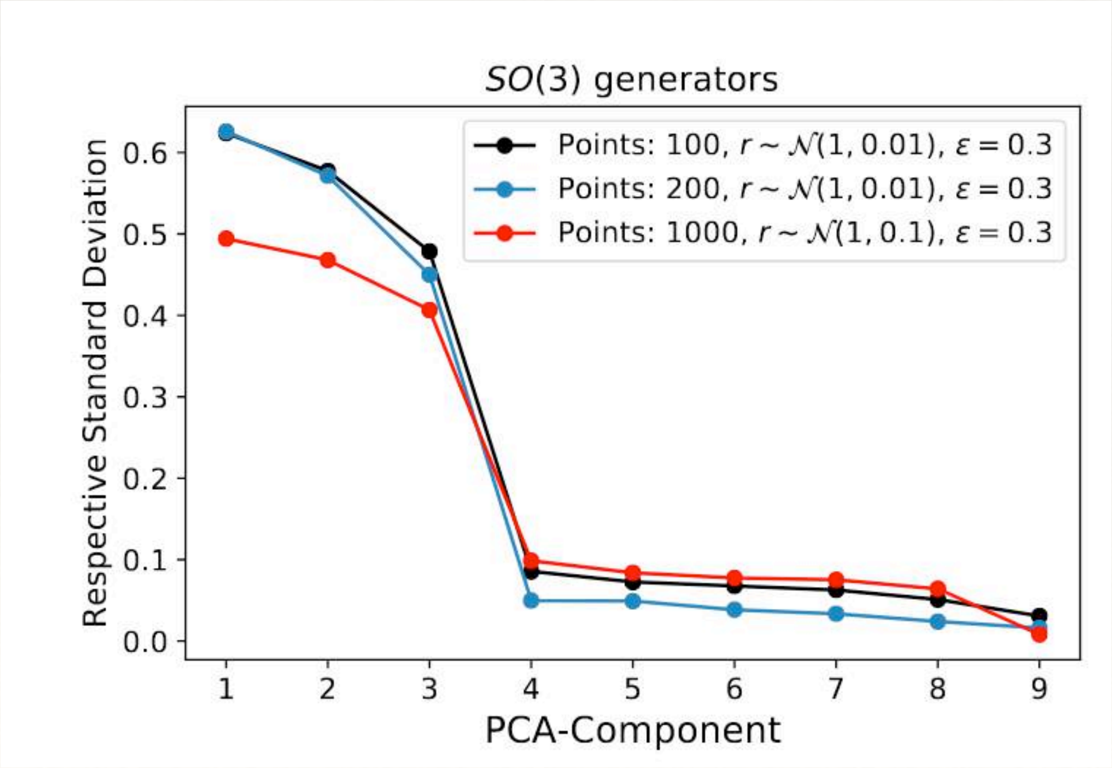
in which ways? and what could we use this for?

Can we teach Machines about Symmetries?



Detecting Symmetries with NNs, 2003.13679
by Sven Krippendorf and Marc Syvaeri

Feeding a symmetric potential to a NN
and assign a multiclass classification task,
get the input of the last hidden layer,
do a dimensional reduction e.g. t-SNE

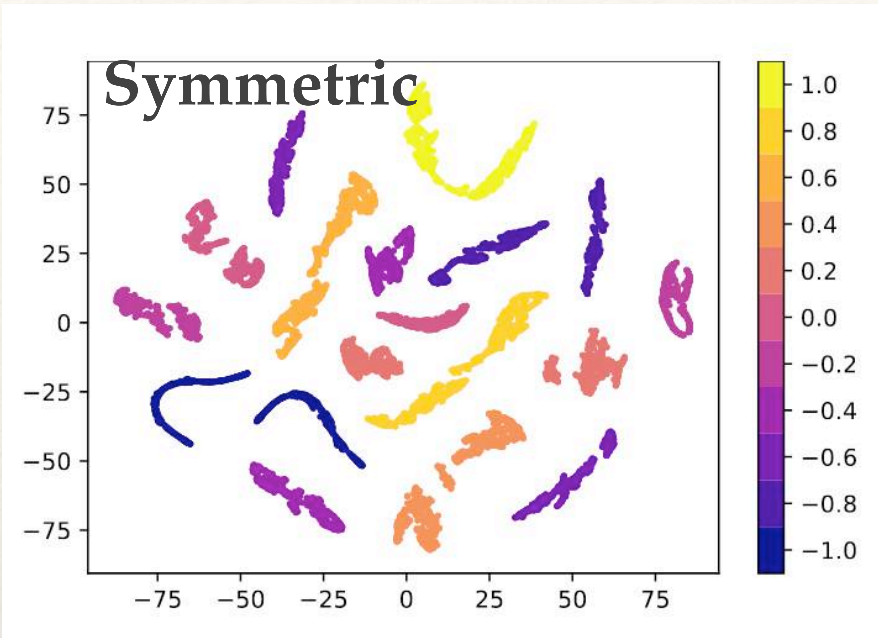


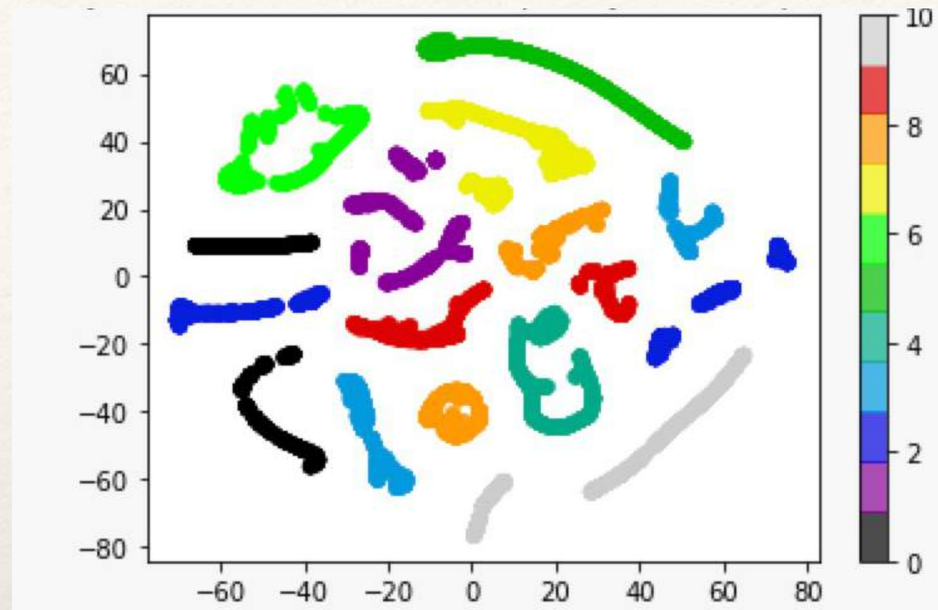
Use the content of this t-SNE
to find the generators of a symmetry
using a regression algorithm

Extremely interesting and useful for
theoretical physics problems
e.g. CY manifolds

But there was a problem:

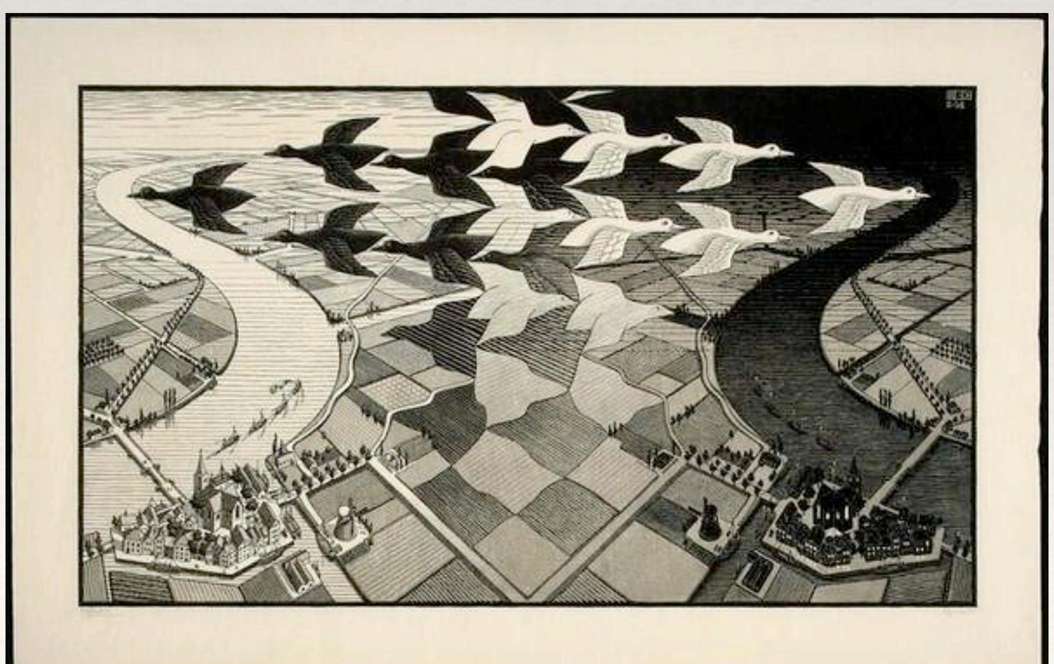
what if the potential had no symmetry, or was only approximate? what if we couldn’t set up a multiclass problem like *Krippendorf & Syvaeri*?





Non-symmetric:
also shows clustering, and replication because that’s what clustering wants to do

---



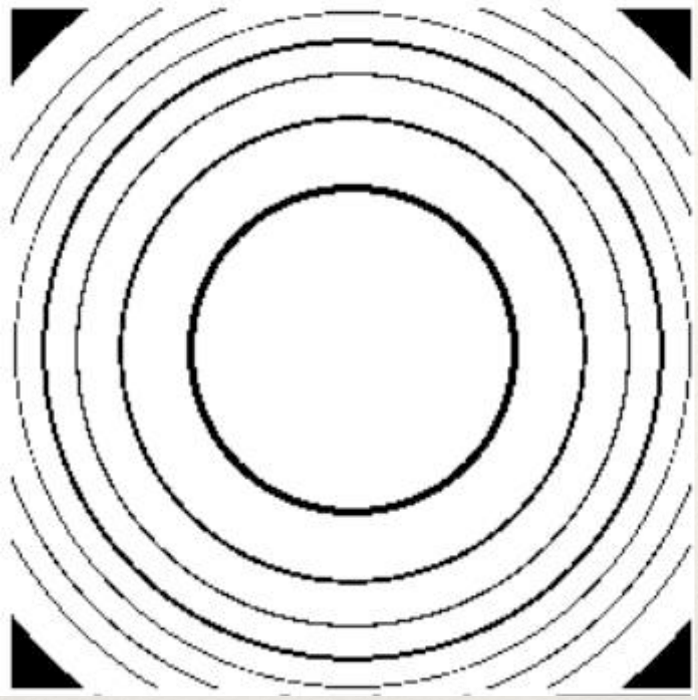
We asked ourselves:

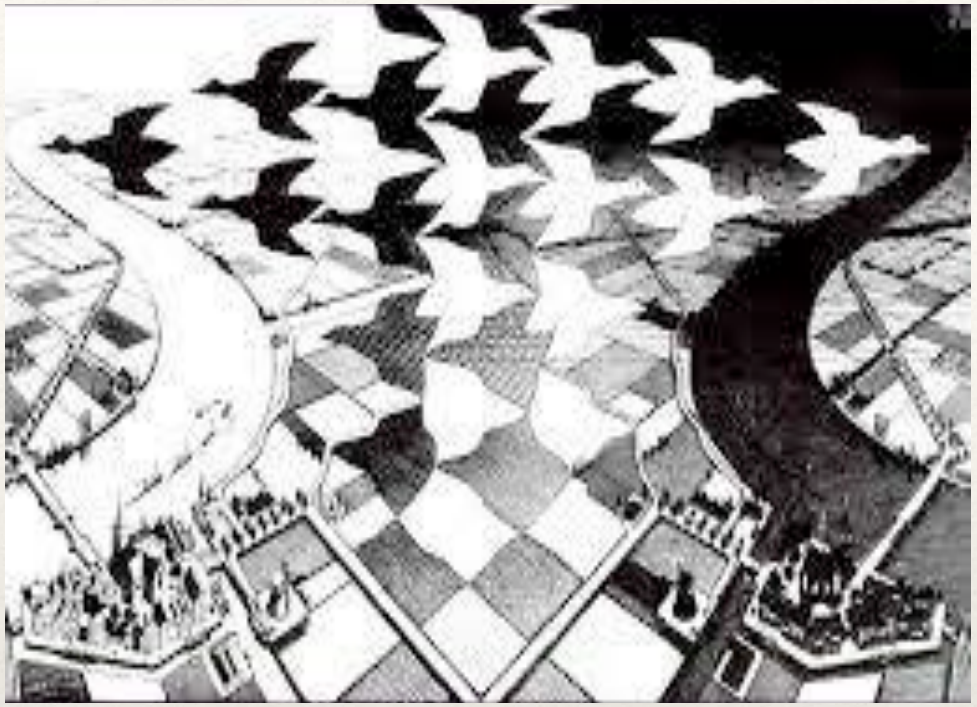
Are there other ways to learn about symmetries which detect no symmetry or approximate levels of symmetry?
and that do not rely on a multiclass task?

We needed a very general procedure

We had to start with something else,
a simpler representation

an image with only two colours





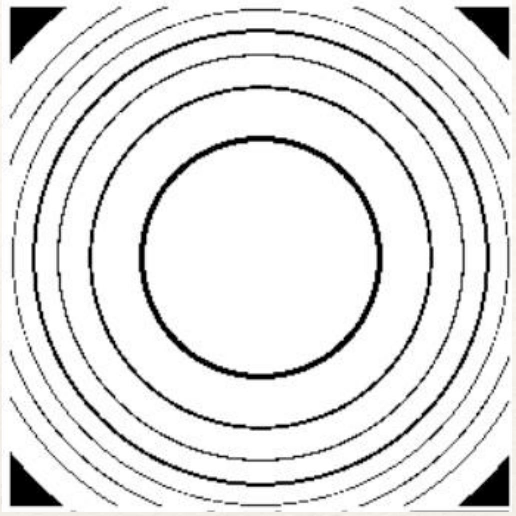
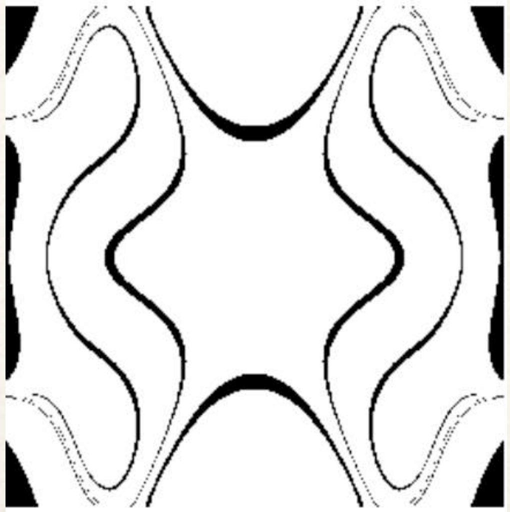
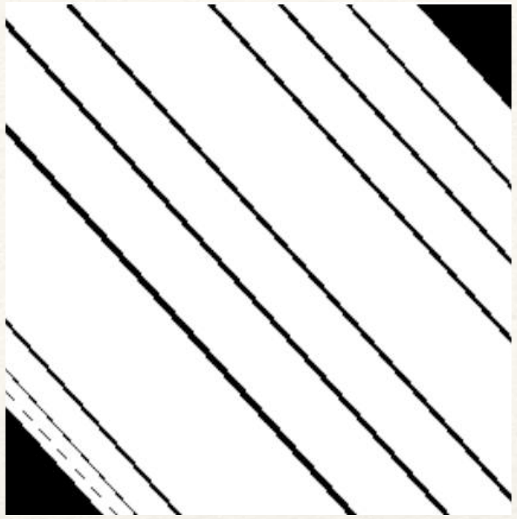
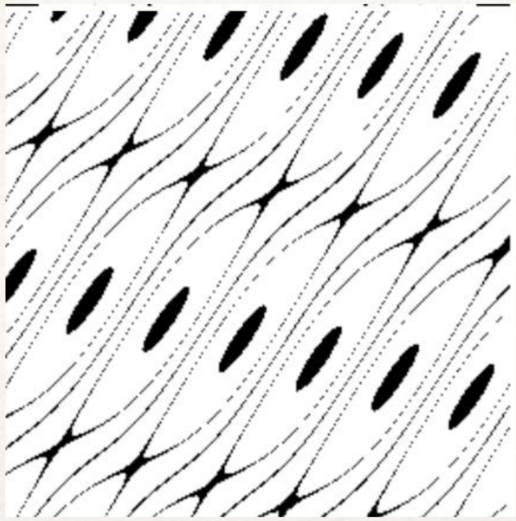
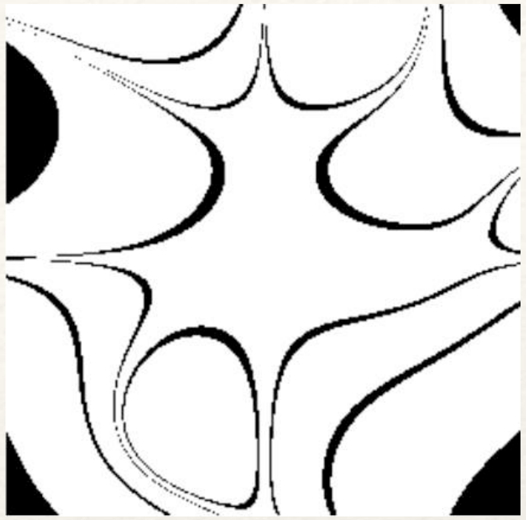
and a universal task:
try to learn as much as possible from this image

dataset = (x , y , 0/1)

and train a FCNN to learn to reproduce the image

then we can ask whether, while learning every detail of the image,
it did realise there was some level of symmetry

To train the FCNN, we build a dataset made of **Physics templates**

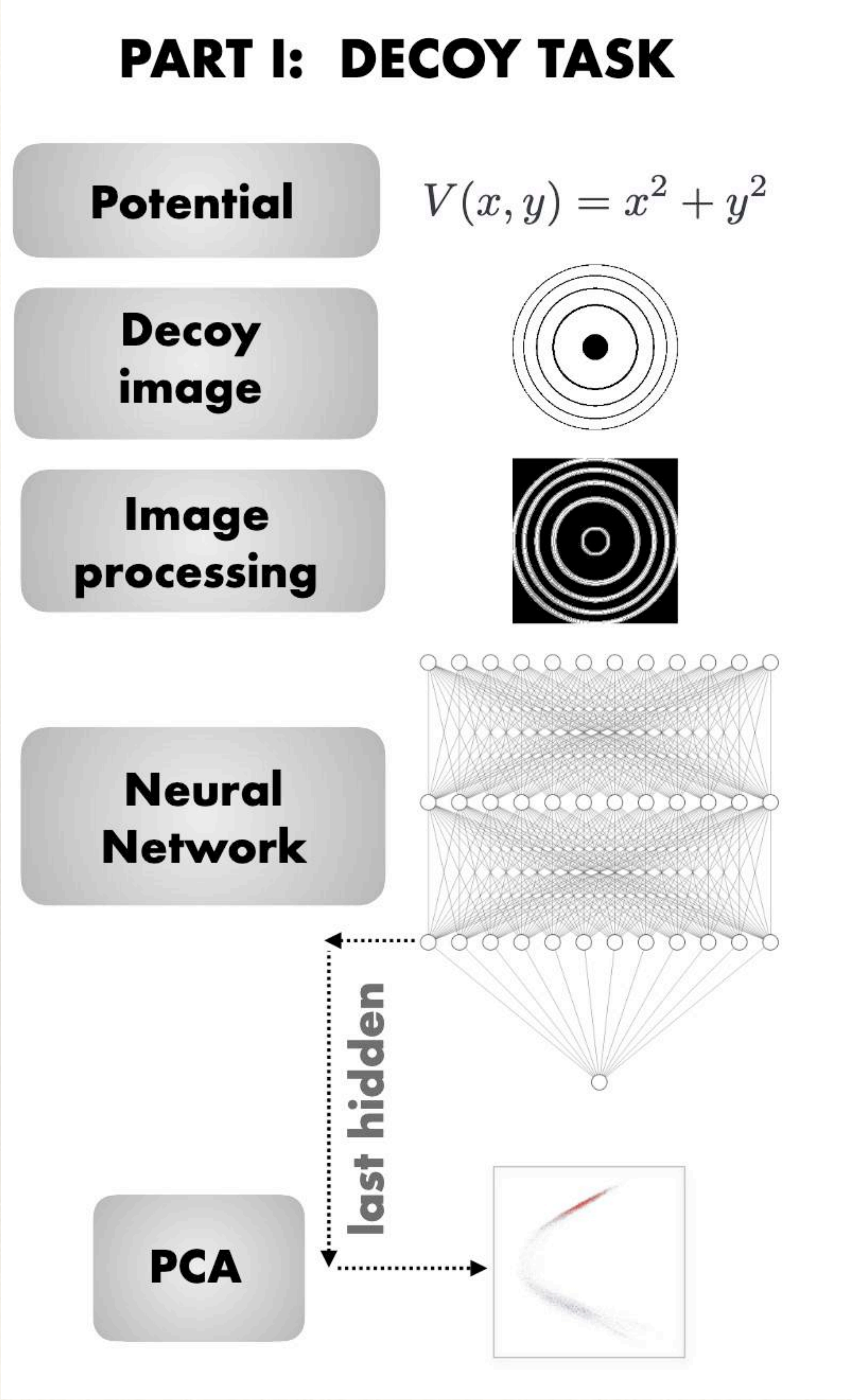
where we know the symmetry properties,
but this information is **not** known to the NN

We pay attention to not **overspecialise** in the physics potentials
FCNN is not allowed to **overfit**, so that it may be more prone to
identify the symmetry

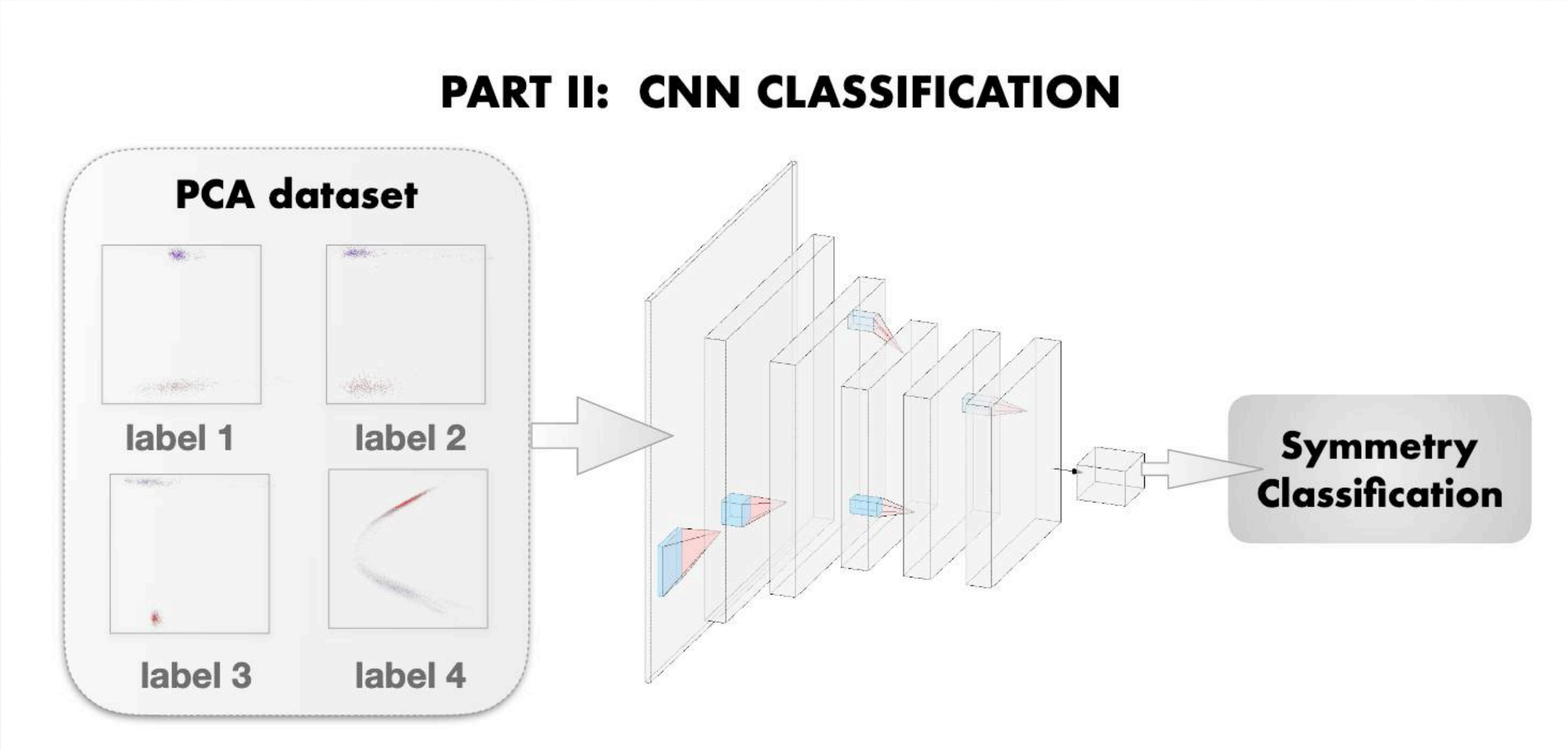
we then get the PCA image from the last hidden layer

at first sight all the PCAs look different, changing
from run to run and from image to image…

Putting it all together



If the FCNN, while paying attention to reproduce the image,
has learnt that there was some symmetry,
the PCAs may encode this learning



We train a CNN, using the PCAs and the physics labels,
to identify symmetries

We find that the PCA- >symm classification does work,
the PCA does contain ***some encoding of the symmetry***

From Physics to Art

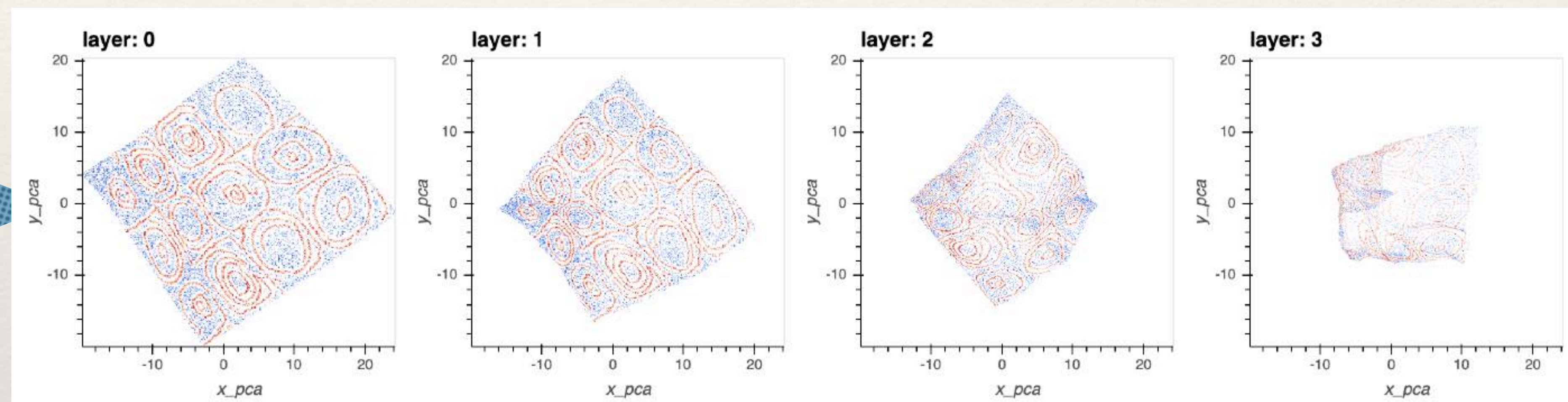
So far have developed an algorithm which takes simple images and produces a **symmetry score**

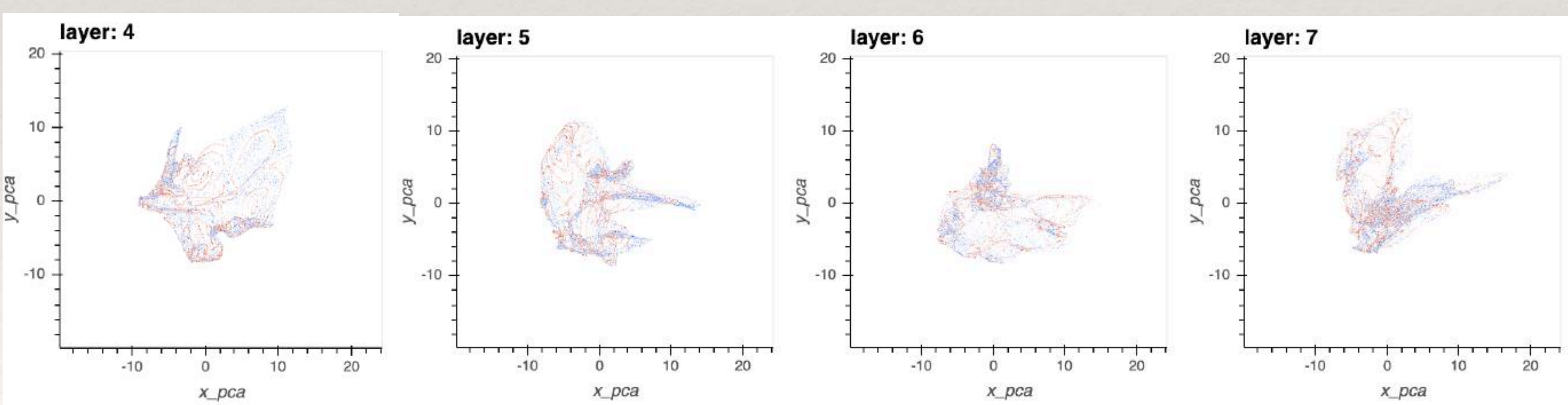
It can be used in any type of 2D images:

a children's sketch, a painting, a photograph…

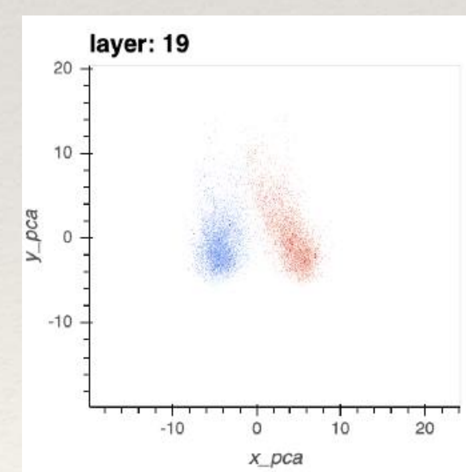
Some examples from Art:

first images-> sketches, then run the algorithm as if they were physics potentials

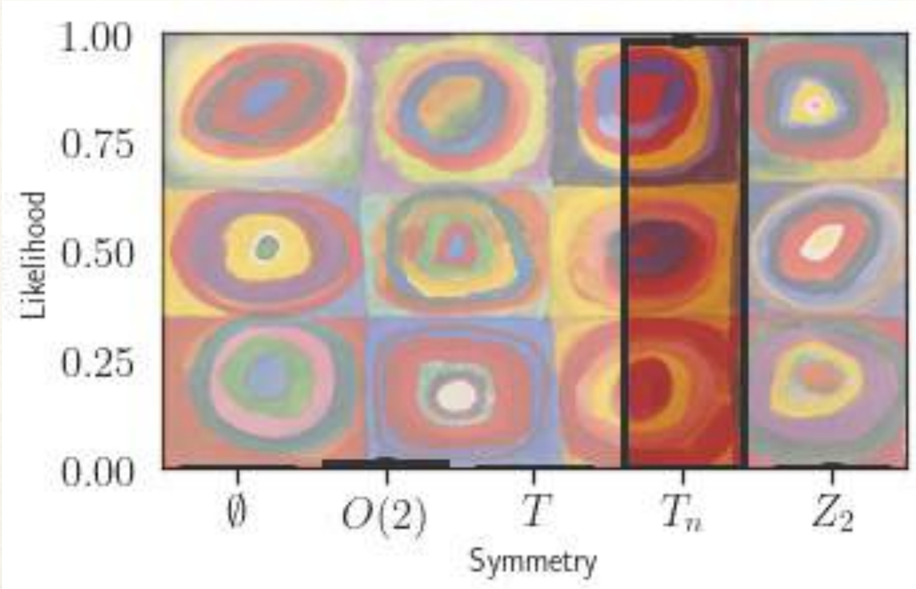


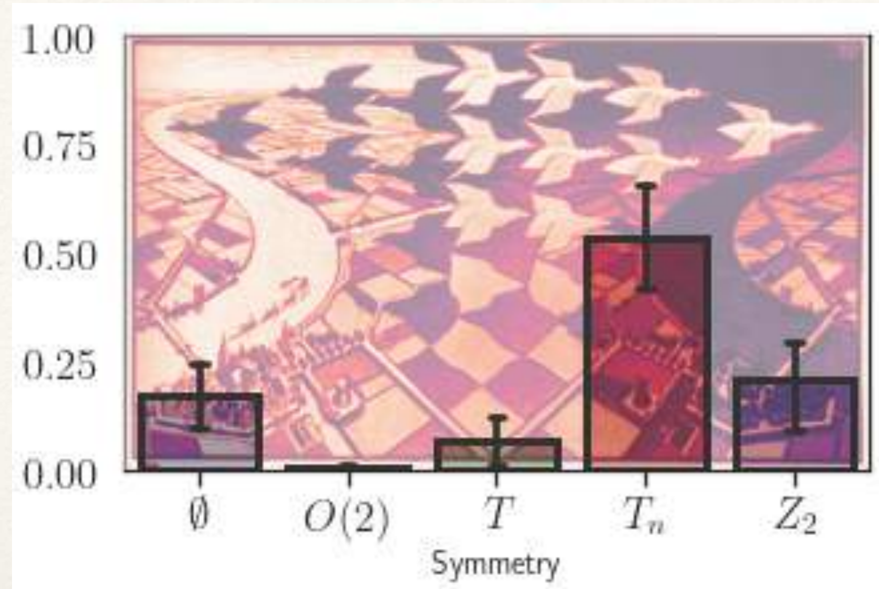


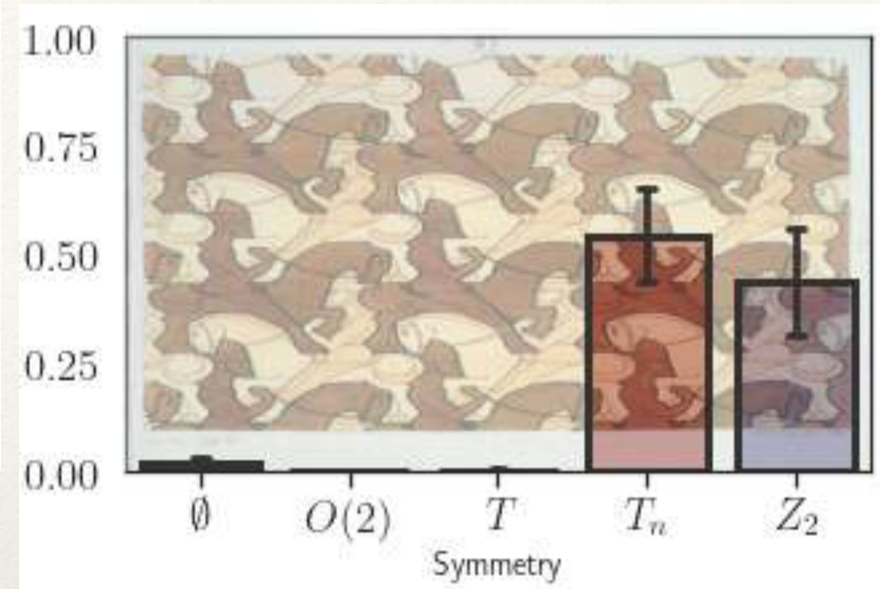
…

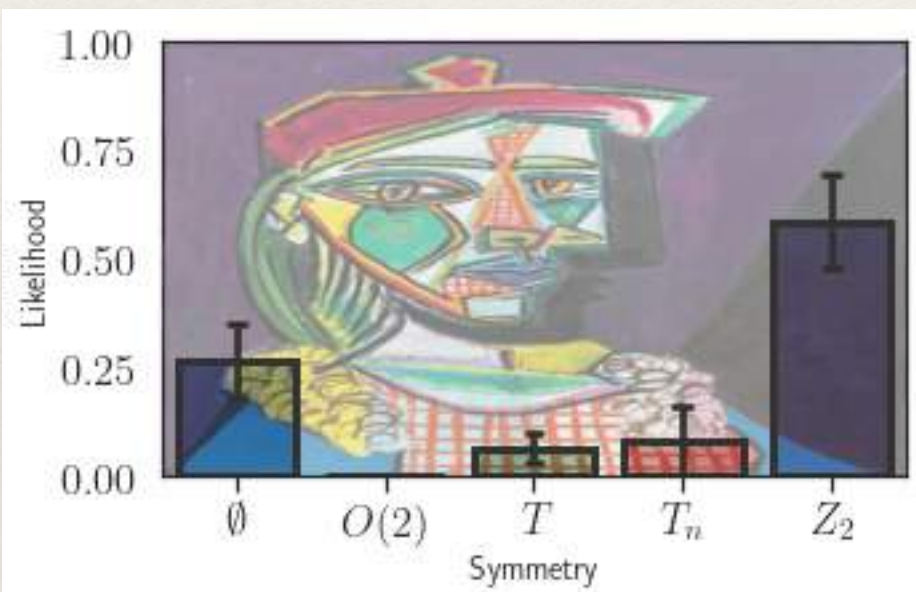


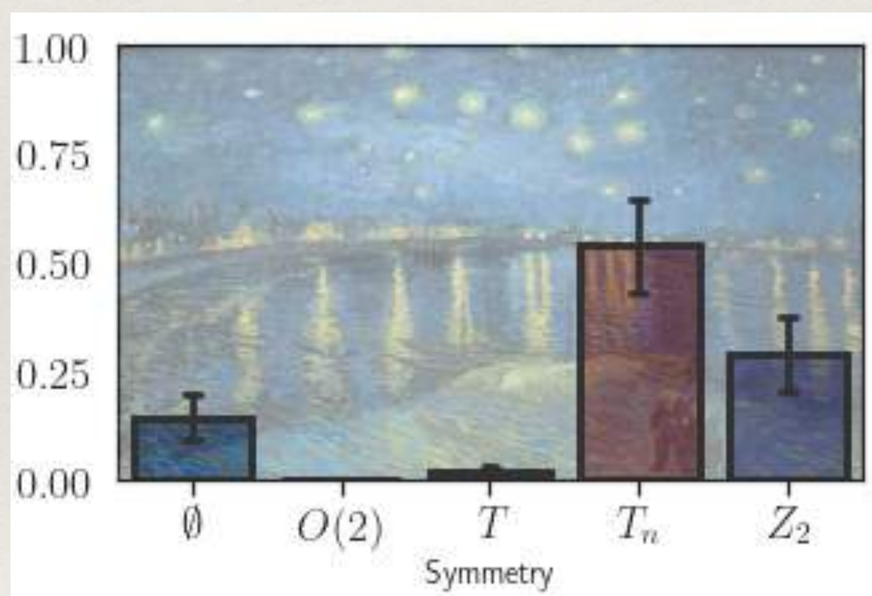
SYMM SCORE?

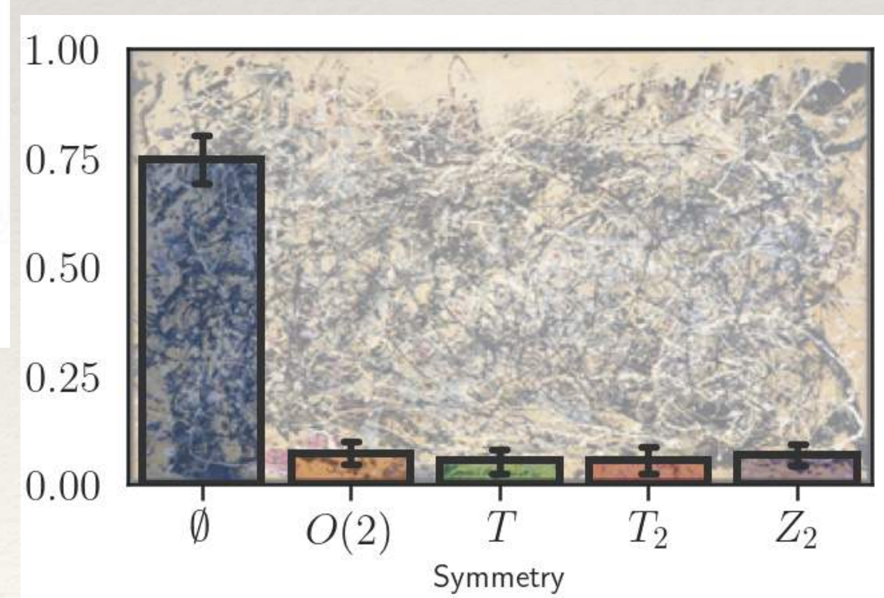












and many more, including children's drawings, fractals, photographs etc

Conclusions and future directions

We are just starting to understand the applications of ML in Physics

They go beyond a mere iteration of our traditional statistical methods: unsupervised methods, generative AI, reinforcement learning…

Through AI methods, there is interesting cross-pollination between our area and others

Yet a very efficient blackbox is not good enough for us, we try to *communicate* with the AI, to find ways to understand its inner workings

Today we learned that an AI does identify and use symmetries, even if only approximate, when *inspecting* an image (decoy task) and this learning can be found in subtle features of the hidden layers

We applied this method to Art, finding that it matches human intuition

<u>What are our next steps?</u>

Images -> sound [Paintings -> Music]

2D -> higher dimensions [Modifications in the decoy task]

Supervised FCNN-> Unsupervised [VAE/GANs]