ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN NEUTRINO PHYSICS

Dr. Saúl Alonso-Monsalve

ETH Zurich

22nd International Workshop on Next Generation Nucleon Decay and Neutrino Detectors (NNN23)

Procida, 13 August 2023

Why do we need machine learning?

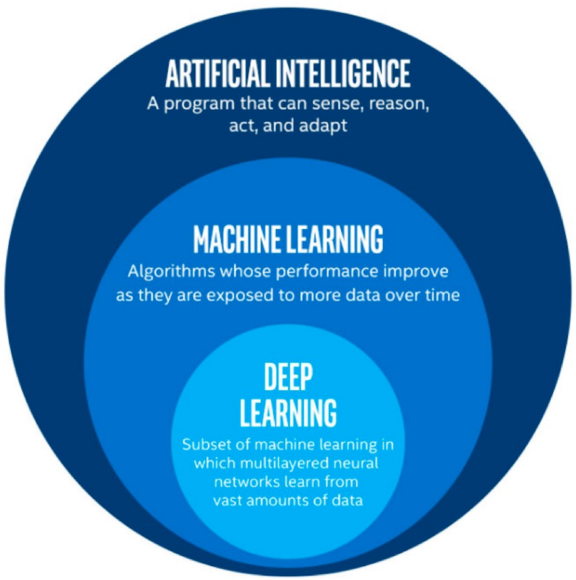
- Neutrinos are elusive particles that rarely interact with matter, making their detection and study incredibly challenging.
  - Neutrino experiments need to collect massive amounts of data in order to see a handful of neutrino events.
  - This data is also very complex, as neutrino interactions can produce a variety of different signatures.

- Challenges for standard algorithms:
  - **Complex pattern recognition**: standard algorithms often struggle to identify intricate patterns in the data.
  - **Potential biases**: human implementation can introduce biases that affect results.

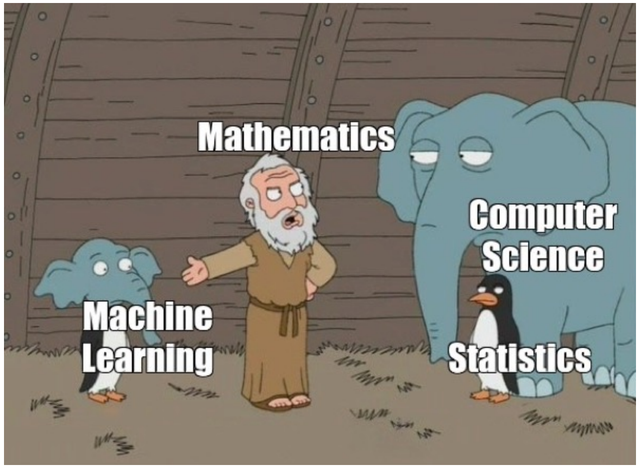
- Machine learning (ML) promise:
  - ML algorithms have the capability to learn from data (simulation) and improve their performance over time.
  - Expect ML to not only meet the challenges but also potentially lead to groundbreaking discoveries in neutrino physics.

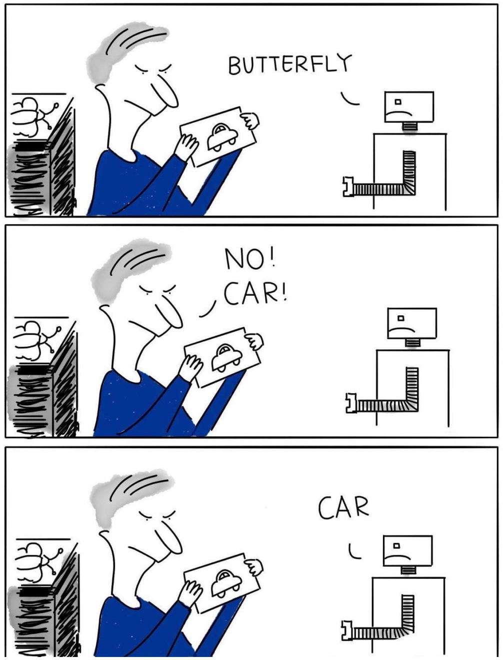
But… what is machine learning?

- ML is a subfield of artificial intelligence (AI).
  - AI: branch of computer science that aims to build algorithms capable of performing tasks typically (traditionally) accomplished using human intelligence.



- ML is learning from data.
  - There is no learning without data.
  - ML algorithms only learn from the data.

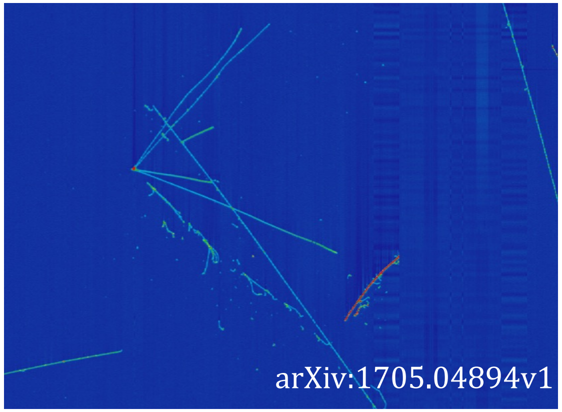


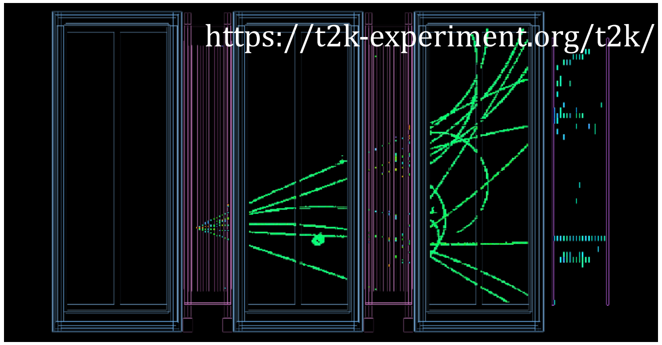


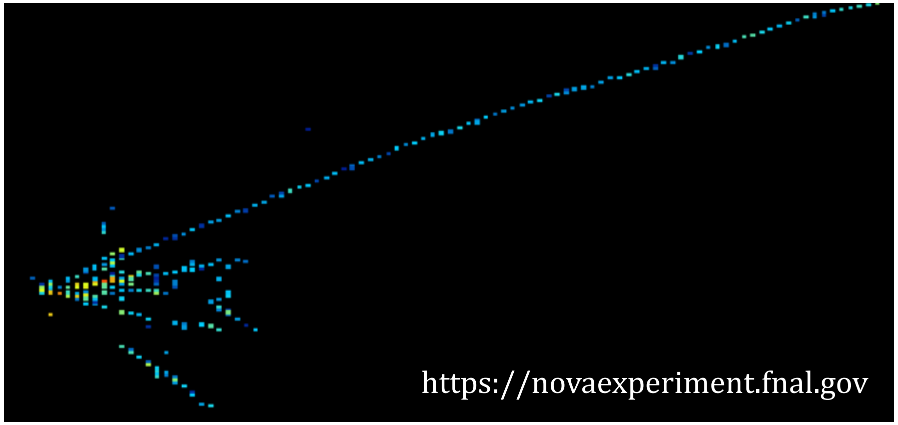
*Source: Stack Exchange

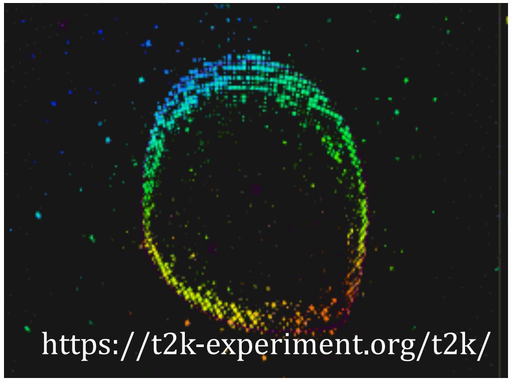
Neutrinos and images

- Neutrino experiments often involve complex detectors that allow us to collect data in the form of images.
  - Typically serves the purpose of detecting and characterizing events generated by neutrino interactions.
  - Examples: LArTPC detectors (DUNE, MicroBooNE, ICARUS), Cherenkov detectors (Icecube, SuperK/HyperK), scintillators (NOvA, T2K/HyperK)…
  - Require powerful algorithms to deal with such complex data.

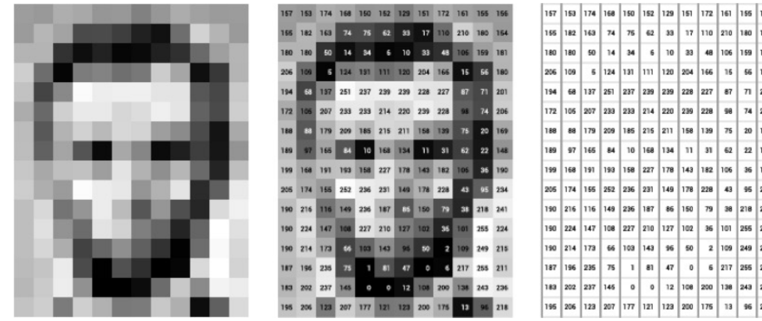








Computer vision



[Source: Openframeworks]

- Study of visual data (images and videos).
  - A computer "sees" a grid of numbers.
  - Massive amount of visual data produced every day.
  - Origin in the late 50s.
- Deep learning has been the dominant approach in computer vision research for the past decade.
  - Applications in many areas: automotive, healthcare, robotics, media, agriculture, security, physics…
- What about MLPs?
  - Fully-connected neural networks (FCNNs), also known as dense neural networks or multi-layer perceptrons (MLPs).
  - Type of artificial neural network where each neuron in one layer is connected to every neuron in the next layer.
  - Require a fixed-size one-dimensional input, resulting in an extremely large number of parameters.
    - In the example, the input (flattened) has 3,072 values. If the first layer has 1,000 neurons, that's 3,072,000 parameters (without bias) for only the first layer!
    - They are not translation invariant.

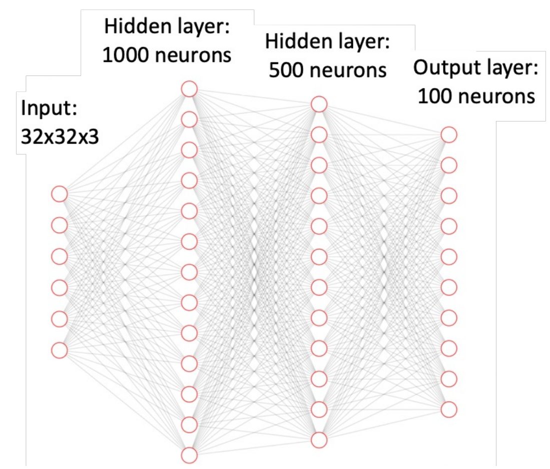
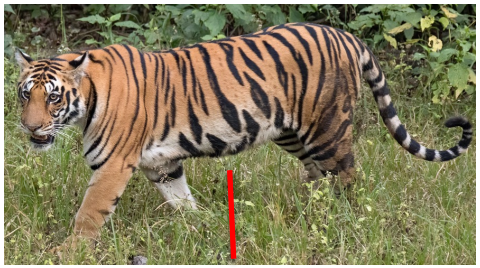
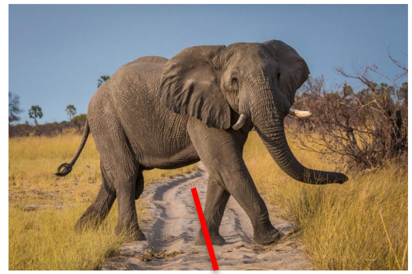


Image recognition (intuition)

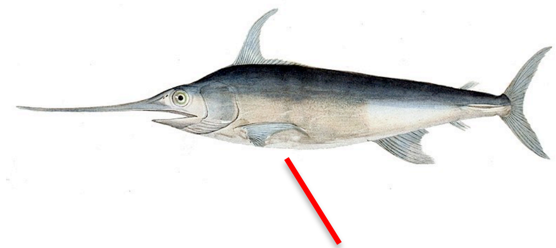
- Humans break down images into different parts before assembling the information back together.



It is orange with black stripes:
it is a tiger.

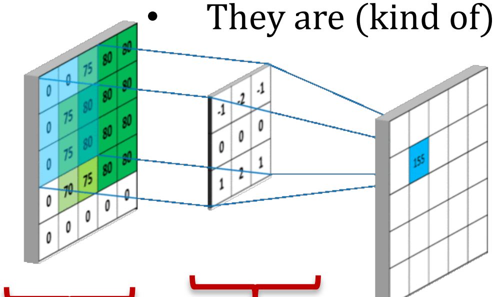


It is gray with a long trunk:
it is an elephant.



It is gray with a long trunk:
it is an elephant.

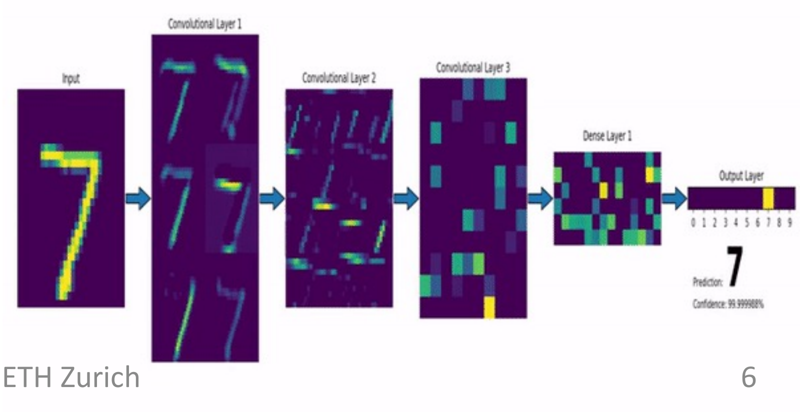
- Convolutional neural networks, or CNNs, are a type of neural network architecture specifically designed for image recognition tasks in computer vision.
  - Convolution: element-wise multiplication and sum of the overlapping elements between the kernel and the input.
  - CNNs use a series of convolutional layers to extract hierarchical features from images.
  - CNNs have achieved state-of-the-art performance in various computer vision tasks.
  - They are (kind of) translation invariant!



input

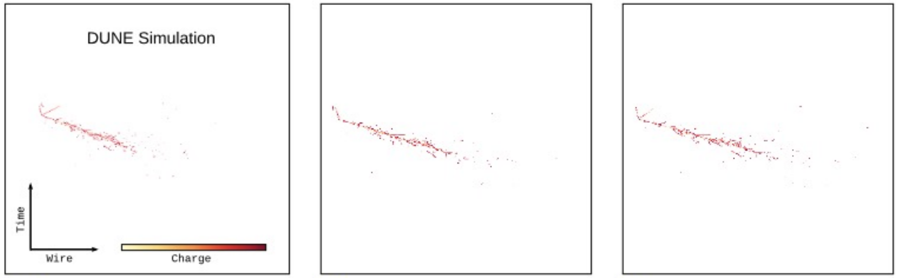
learned
kernel

output
(feature map)

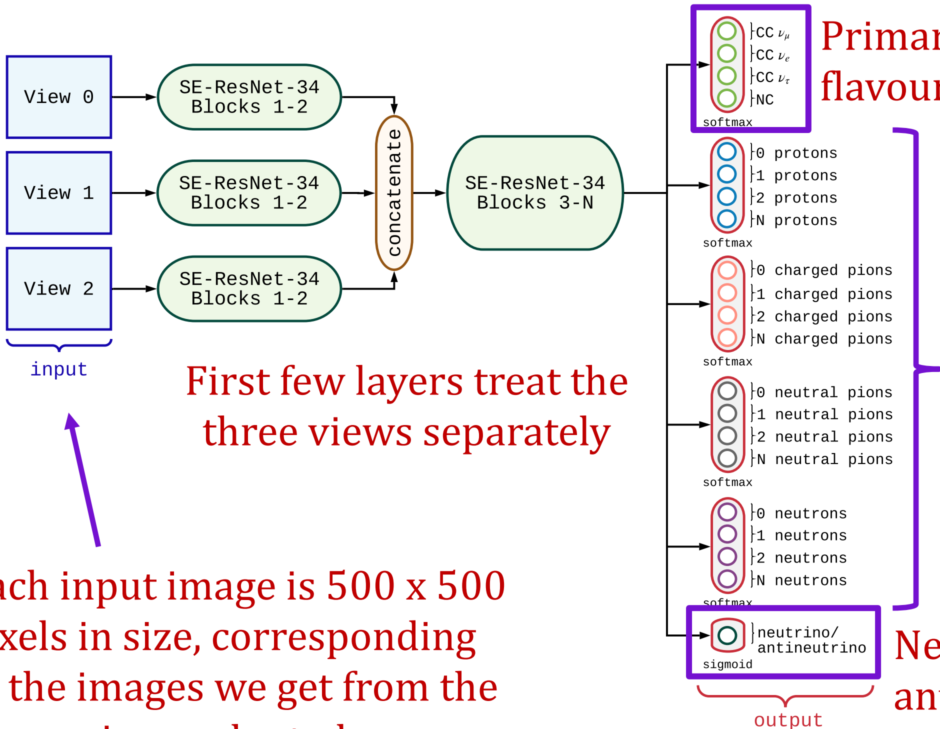


CNN example in neutrino physics: the DUNE CVN

- Inspired by the NOvA CVN: J. Phys.: Conf. Ser. 888 012063.
- CVN stands for "Convolutional Visual Network".
  - Publication: Phys. Rev. D 102, 092003.
  - The DUNE FD provide three "images" of each neutrino interaction.



(a) View 0: induction plane (U) (b) View 1: induction plane (V) (c) View 2: collection plane (Y)



Primary output: flavour identification

Secondary outputs: particle counting for exclusive final states

Neutrino / antineutrino

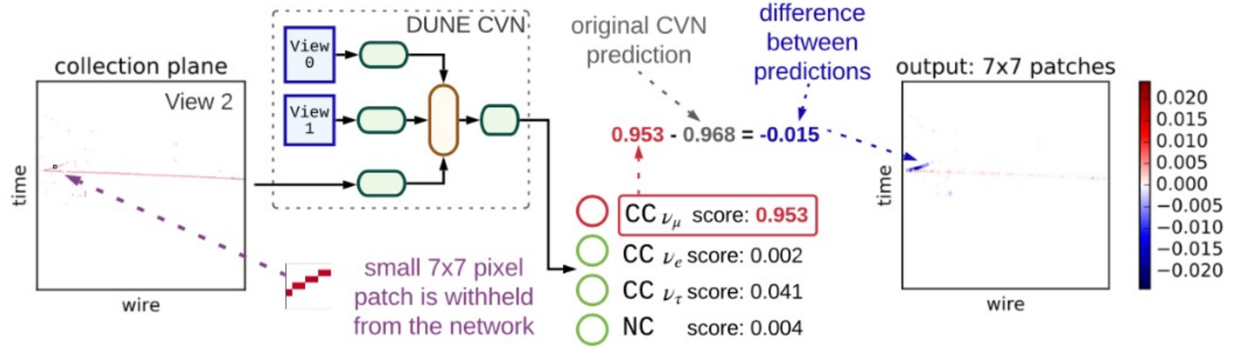
First few layers treat the three views separately

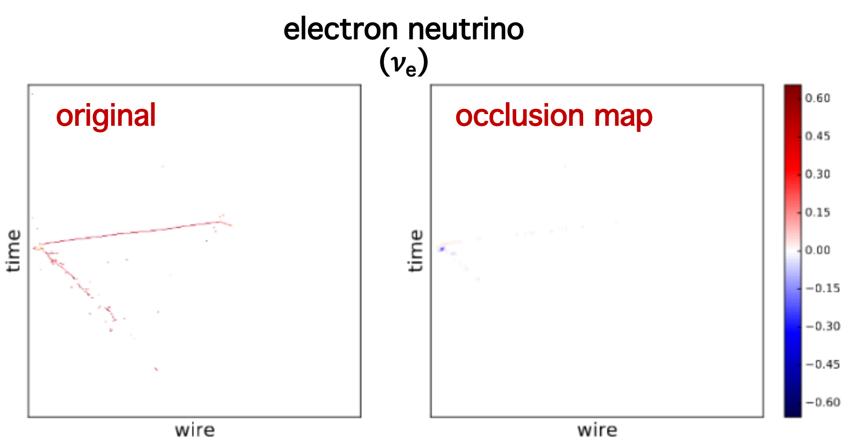
Each input image is 500 x 500 pixels in size, corresponding to the images we get from the three wire readout planes.

- The **primary output results** (flavour) were **used in the official DUNE neutrino oscillation sensitivity analyses.**
  - DUNE Technical Design Report (TDR, 2020): arXiv:2002.03005.
  - DUNE Long-baseline (LBL) analysis (2020): https://doi.org/EPJC/S10052-020-08456-Z.
  - Milestone for the experiment!

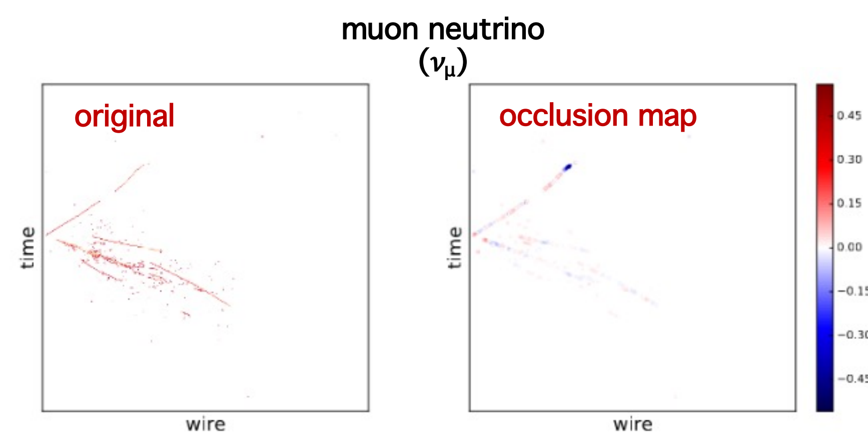
Understanding the CVN

- Occlusion tests:
  - Hide parts of the images and check how the CVN reacts to the changes.





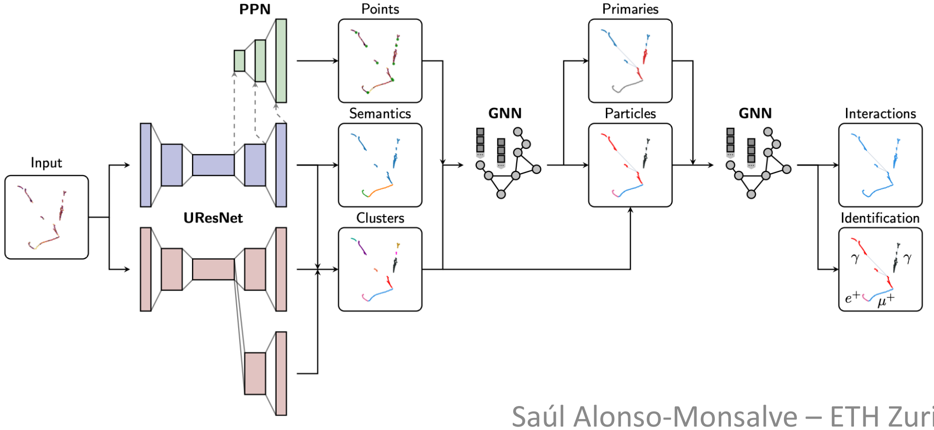
Removing the start of the electron shower reduces the CC ν_e score, as expected



The CVN finds the vertex a bit ambiguous, but it uses the end point of the muon to gain a handle on the event type.

Going further: handling sparse data

- In particle physics and astrophysics, data is often sparse due to the nature of the objects being studied or the particles detected.
- This poses a challenge for computer vision, as standard CNNs are designed to work with dense data. To address this, researchers are developing new algorithms and techniques specifically tailored to sparse data.
  - For example, one approach is to use Submanifold Sparse Convolutional Networks (SSCN), where the convolution operation is performed only on the non-zero elements of the sparse data, resulting in an efficient computation.
  - Example: "Scalable, End-to-End, Deep-Learning-Based Data Reconstruction Chain for Particle Imaging Detector" (arXiv:2102.01033)



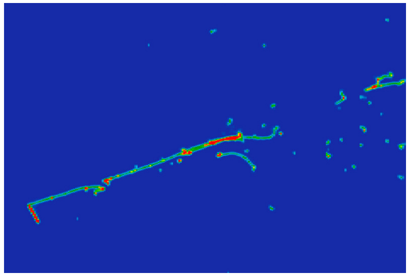
"Dense" image

[Source: britannica.com]

All pixels might be helpful for the classification.

Ideal for standard CNNs.

"Sparse" image



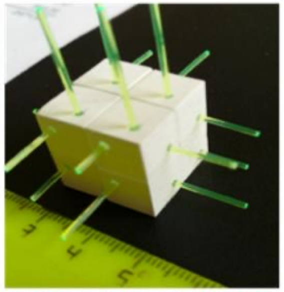
Most pixels are background.

A standard CNN would perform loads of useless computations.

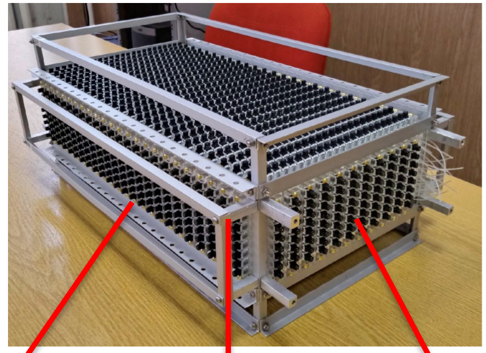
Another solution is to use Graph Neural Networks (see Adam's talk!)

Our current work!

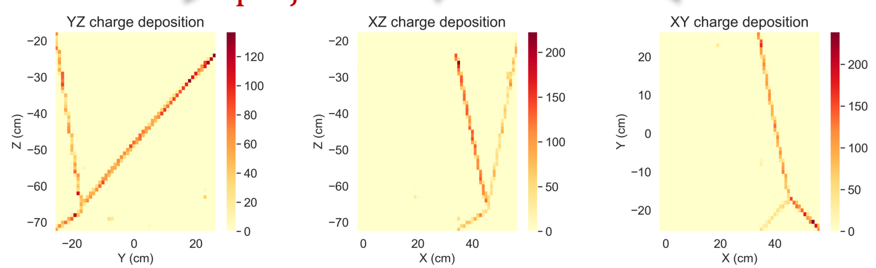
- Case study: a detector concept analogous to the SuperFGD detector from the T2K experiment.
  - Part of the upgrade of the near detector (ND280) of the T2K experiment in Japan.
  - Fully-active fine-grained scintillator (FGD) with three views.
  - 2M optically independent cubes, 1 cm^3 per cube.
  - Spatial localisation of scintillation light.



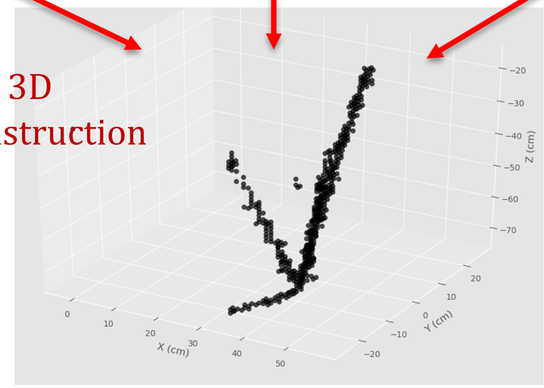
JINST 13 (2018) 02, P02006
NIM A936 (2019) 136-138



2D projections



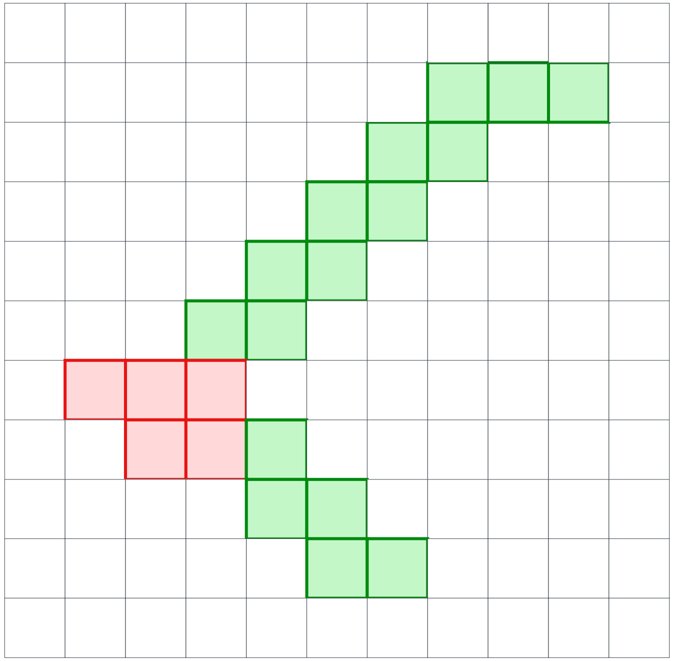
3D reconstruction



- Goal: develop a deep-learning-based analysis strategy that does not depend on the neutrino interaction model:
  a) Distinguish between single and multi-primary-particle hits.
  b) Fit the trajectory of single-particle objects.
  c) Understand the activity at the vertex of neutrino interactions.

Approach

a) **Hit identification.**
b) Particle trajectory fitting.
c) Vertex activity fitting.



Hit identification: single vs multi-particle hits

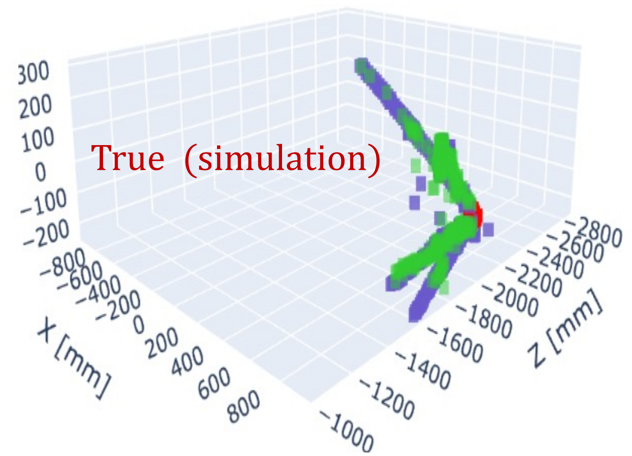
- Classify each individual hit as:
  - **Single-particle hit**: only one particle passes through the hit cube.
  - **Multiple-particle hit**: at least two different particles pass through the hit cube.
  - **Other**: crosstalk or ghost.

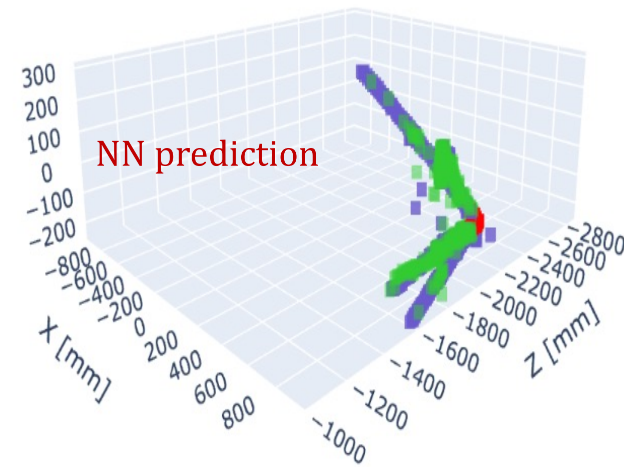
- Using a submanifold sparse U-Net-based neural network architecture (https://arxiv.org/abs/1706.01307).
  - More computationally efficient than standard CNNs.

- Efficiencies:

	True multi-particle	True single-particle	True other
Pred. multi-particle	**0.7777**	0.1511	0.0711
Pred. single-particle	0.0055	**0.9654**	0.0291
Pred. other	0.0079	0.0479	**0.9442**

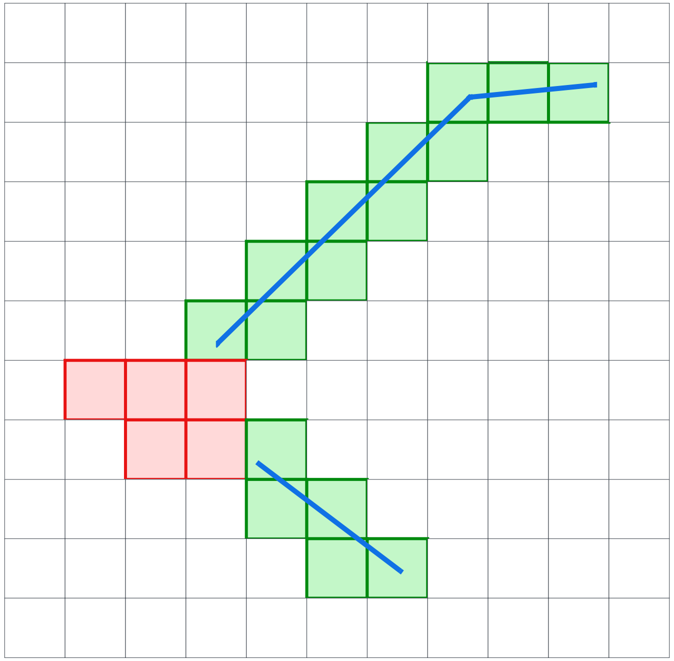
- Excellent single-particle isolation accuracy allows running a further NN-based track trajectory fitting on single particles, relying on detailed MC simulations of single particles.





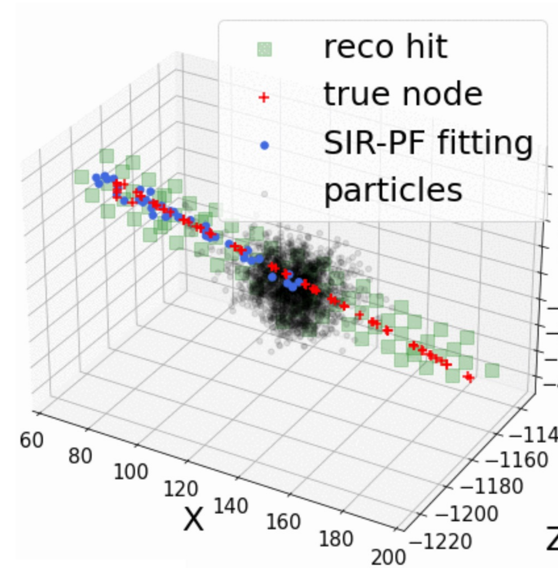
Approach

a) Hit identification.
b) Particle trajectory fitting.
c) Vertex activity fitting.

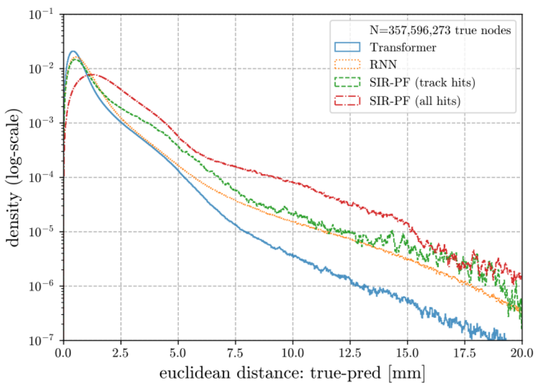


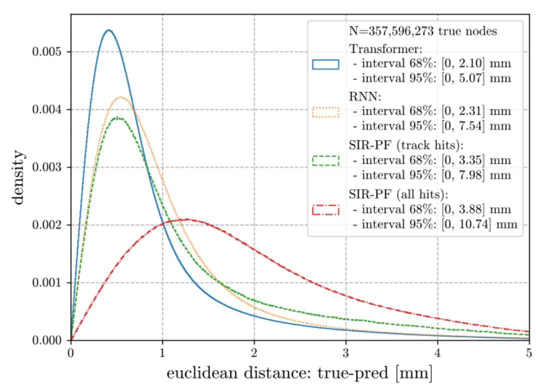
Fitting particle trajectories

- Each algorithm outputs the fitted 3D trajectory point for each input hit.
  - Sampling Importance Resampling Particle Filter (SIR-PF): first reconstructed hit used as prior. The likelihood relies on a precomputed 5-dimensional histogram.
  - Recurrent neural network (RNN): five bi-directional GRU layers, 50 hidden units.
  - Transformer: 5 encoder layers, 8 heads, hidden size of 64.

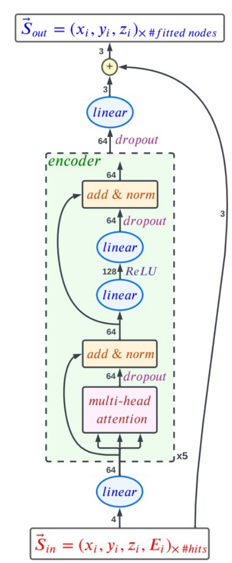


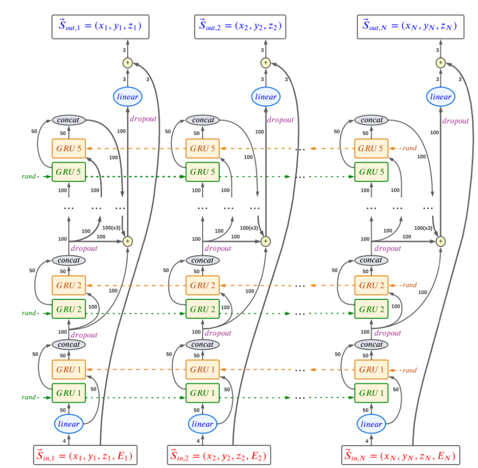
- Main results ([Commun. Phys 6, 119 (2023)](#)):





>30% better transformer resolution compared to the SIR-PF

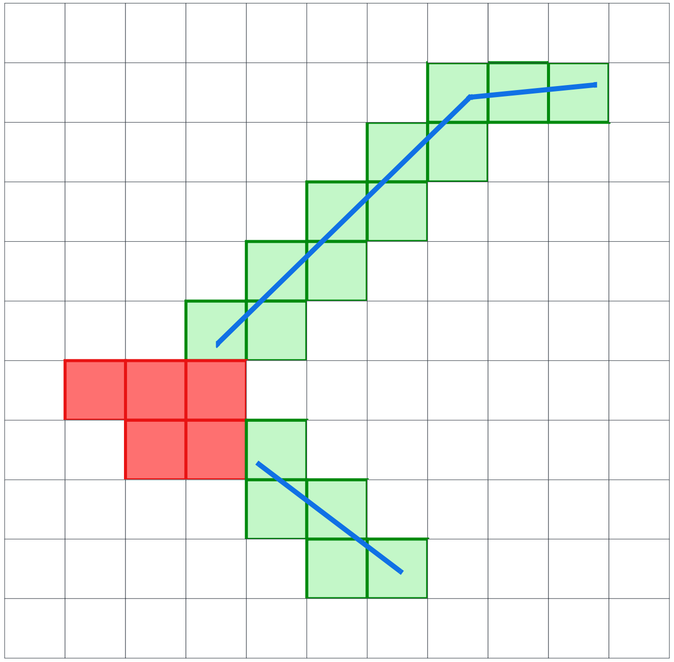




- The improved trajectory fitting significantly improves the charge reconstruction, PID by range, and momentum by curvature.

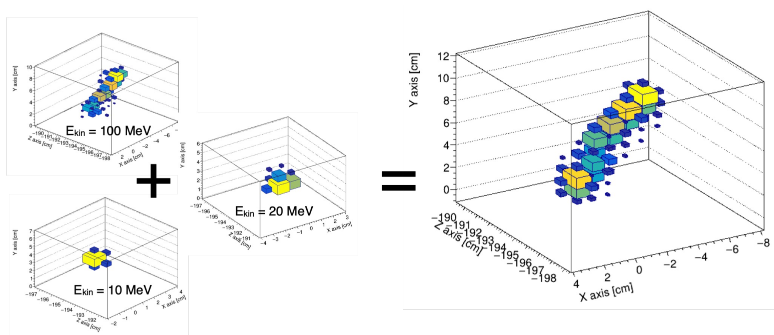
Approach

a) Hit identification.
b) Particle trajectory fitting.
c) **Vertex activity fitting.**



Vertex activity: standard fitting method

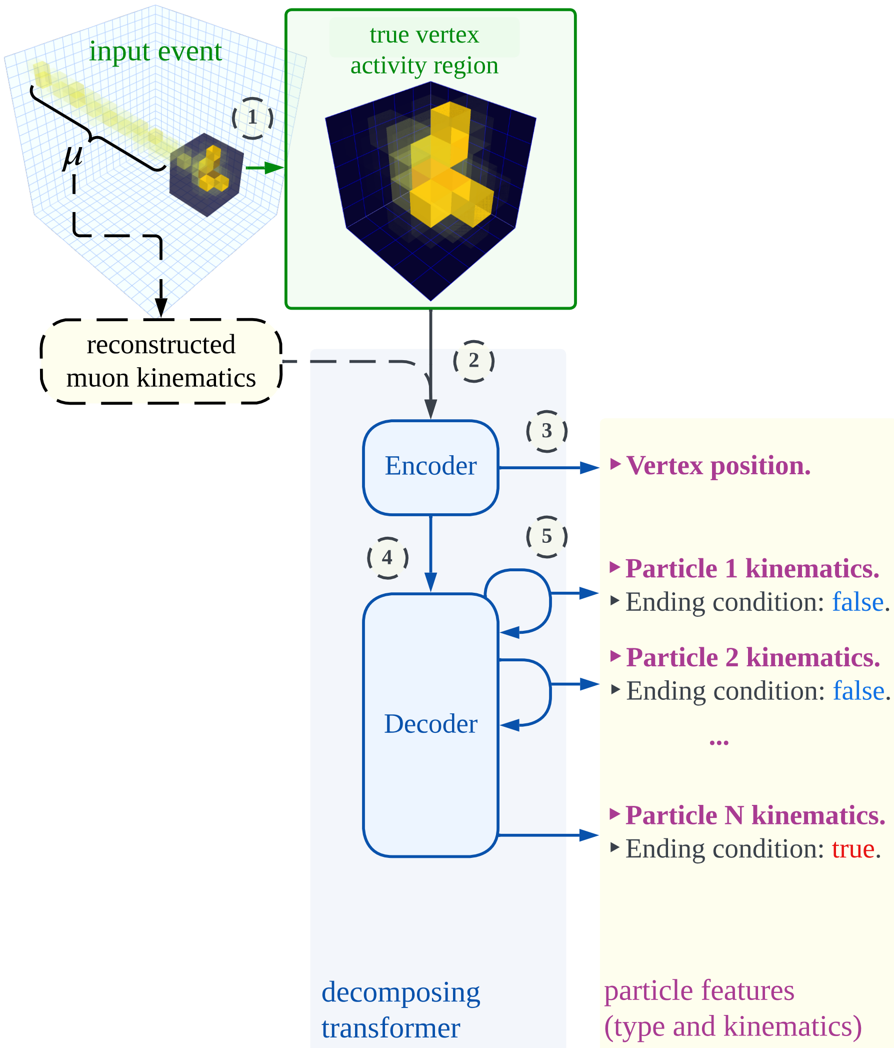
- Vertex activity (VA): particles releasing their energy in the proximity of the neutrino interaction vertex but that do not show visible tracks from where the kinematics can be reconstructed.
  - A "blob" of scintillation light is observed.



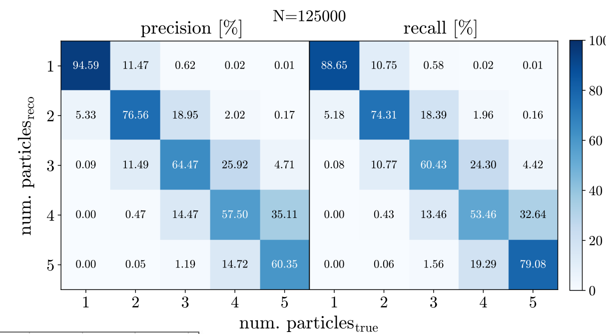
- Standard VA fitting method:
  - ▪Goal: build the neutrino VA in forward folding from the sum of single particle reconstructed objects.
    - Particle information to reconstruct: # of particles (mostly protons), energy, direction, vertex position.
  - ▪Method: likelihood fitting.
    1. Simulating any possible combination of the VA parameters and build VA.
    2. Finding the VA 3D image (e.g. SFGD hits) that "best fits" the data and find the "best-fit" parameters.

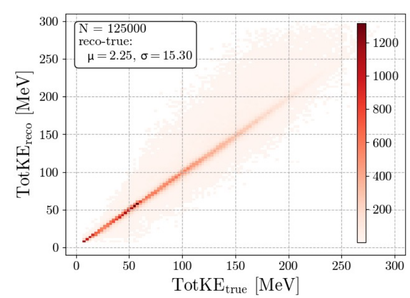
- The fitting method is highly computationally expensive.
  - Requires a large number of combinations of parameters to be simulated.
  - Unfeasible in practice.

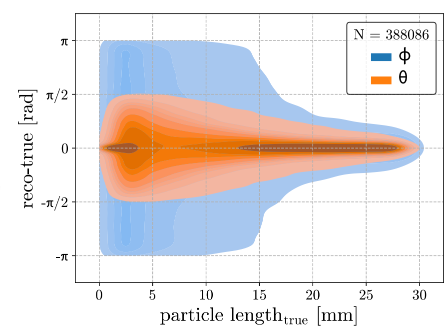
Alternative: deep-learning approach



- **Event config: 1 muon, 1-5 protons.**
  - 70% accuracy in reconstructing the correct number of particles.
    - >98% assuming a ±1 error.
  - ~2 mm vertex resolution.
  - Good reconstruction of kinematics.

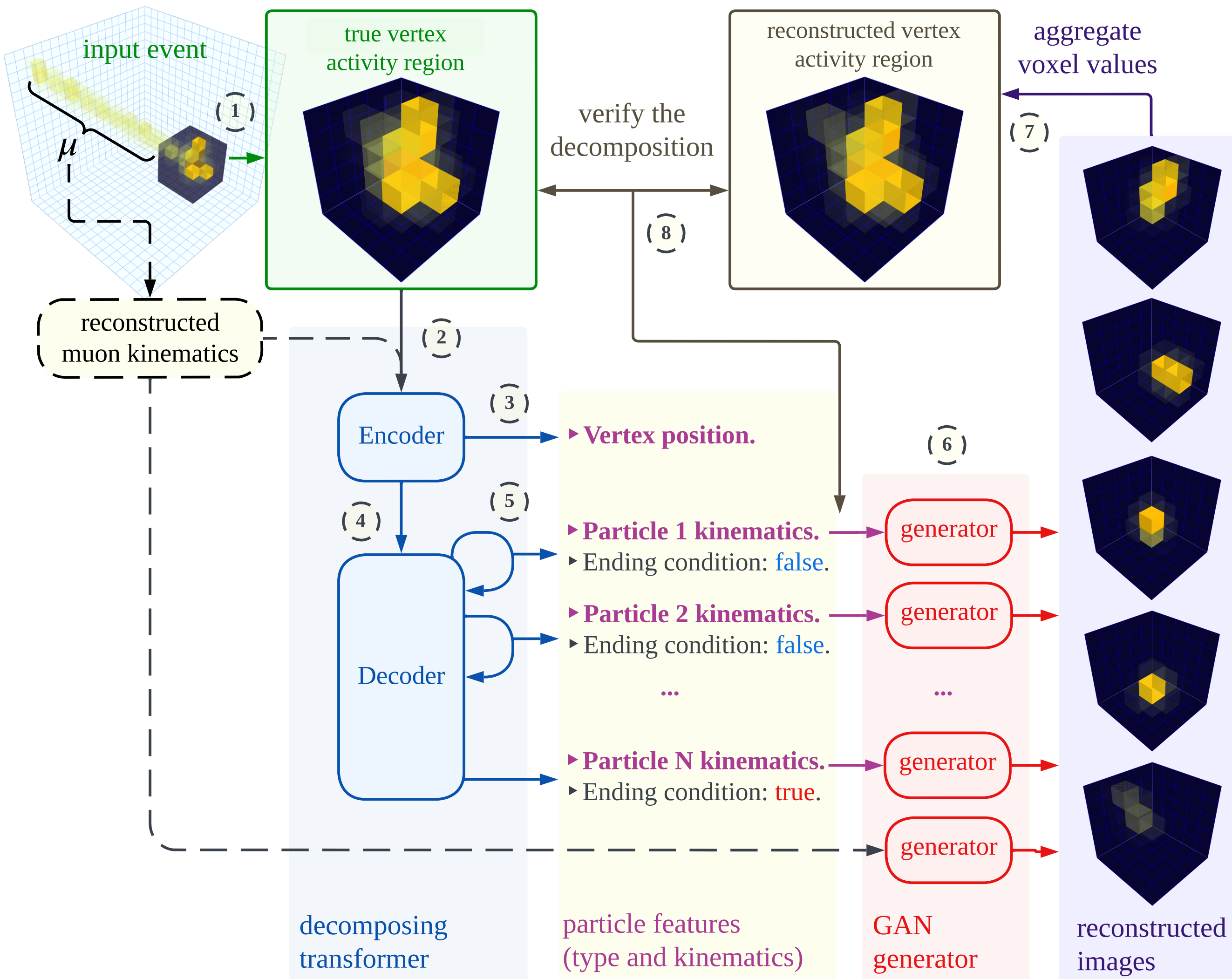




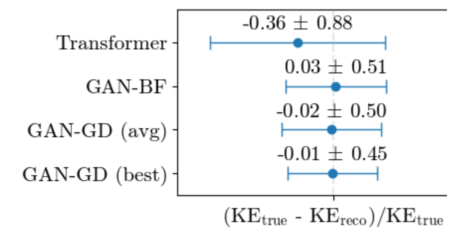


- **Also tested configurations with nuclear clusters (deuterium, tritium)!**

Alternative: deep-learning approach



- The generator (GAN) is fully differentiable.
  - Run a gradient-descent minimiser to optimise the particle kinematics further.



- *GAN-BF = Brute force.*
- *GAN-GD (avg) = Gradient-descent (average result).*
- *GAN-GD (best) = Gradient-descent (best result).*

Summary

- Machine learning could be a key tool for neutrino experiments.
  - In particular, for detectors that provide fine interaction details but are hard to analyse using traditional methods.
- Strongly linked with computer vision.
  - Most of our data is in the form of images.
- Successful application to different problems, such as:
  - Flavour identification, final-state particle counting, track fitting, etc.
- Current and future work requires an extensive validation of the methods and application to experimental data.

Bonus

- Asked the Bing AI model (https://www.bing.com/images/create) to generate images from the following prompt:

  "*Neutrino interaction powered by AI happening at Procida Island in Italy*".

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN NEUTRINO PHYSICS

Dr. Saúl Alonso-Monsalve

ETH Zurich

22nd International Workshop on Next Generation Nucleon Decay and Neutrino Detectors (NNN23)

Procida, 13 August 2023