Reconstruction of electromagnetic showers in calorimeters using Deep Learning

Polina Simkina

Fabrice Couderc, Julie Malclès, Mehmet Özgür Şahin

CEA IRFU

IN2P3/IRFU Machine Learning Workshop
27/09/2022

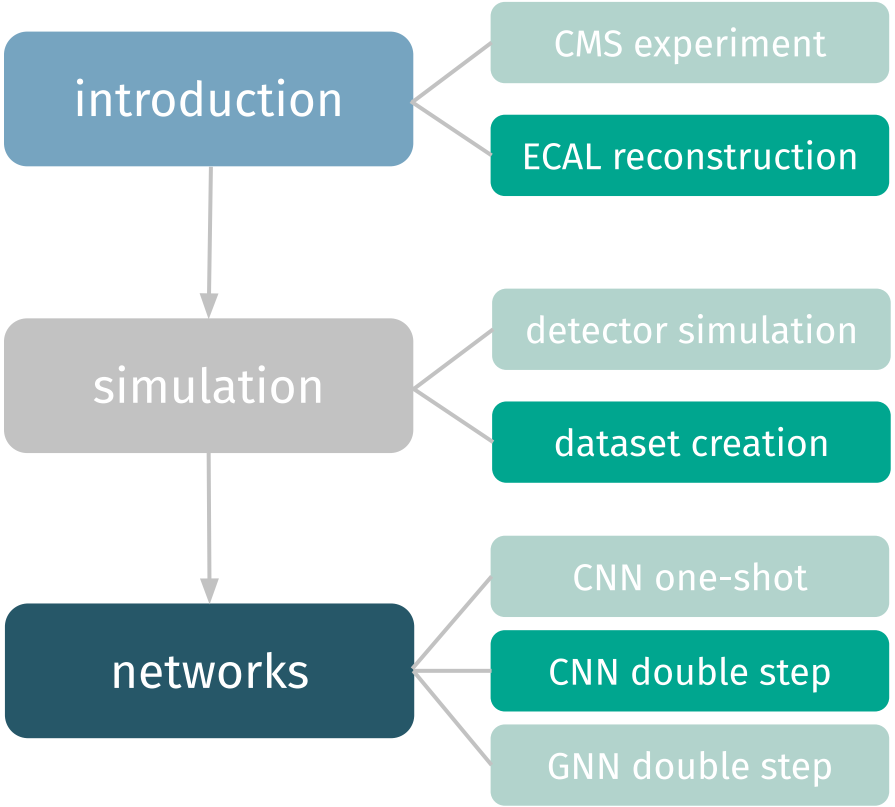
polina.simkina@cern.ch

Outline

Novel Machine Learning method for calorimeter reconstruction.

Multiple architectures (based on CNN, GNN) developed and tested.

Focusing on the reasoning behind using particular methods.



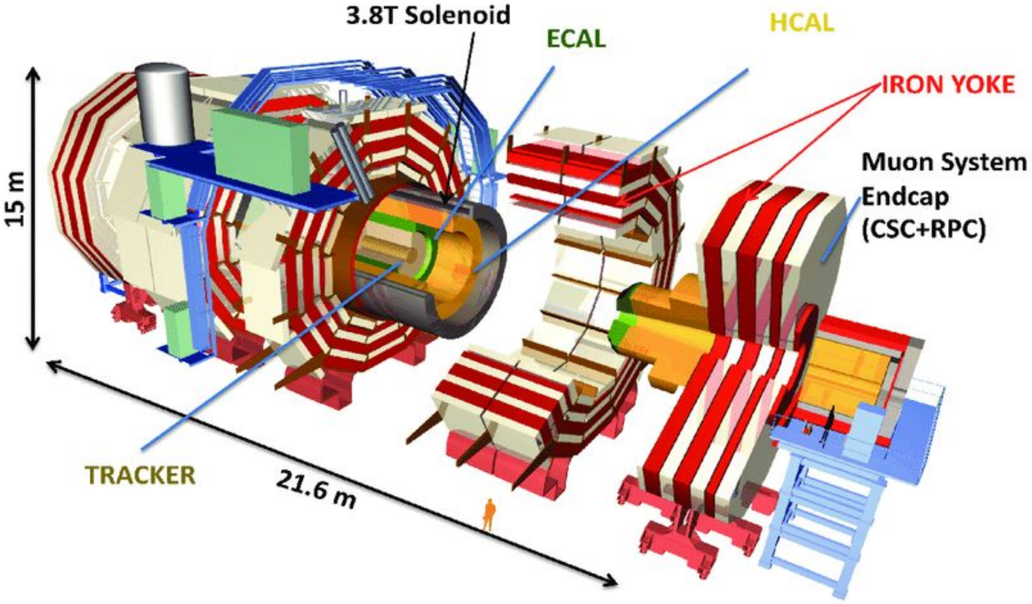
Introduction

CMS experiment

Discovery of the **Higgs boson** in 2012 (along with ATLAS).

Physics scope: probe **standard model** and search for **physics beyond standard model**.

Uses **proton-proton collisions** at the center of mass energy from 7 TeV to 13.6 TeV.



Electromagnetic CALorimeter

Homogeneous calorimeter.

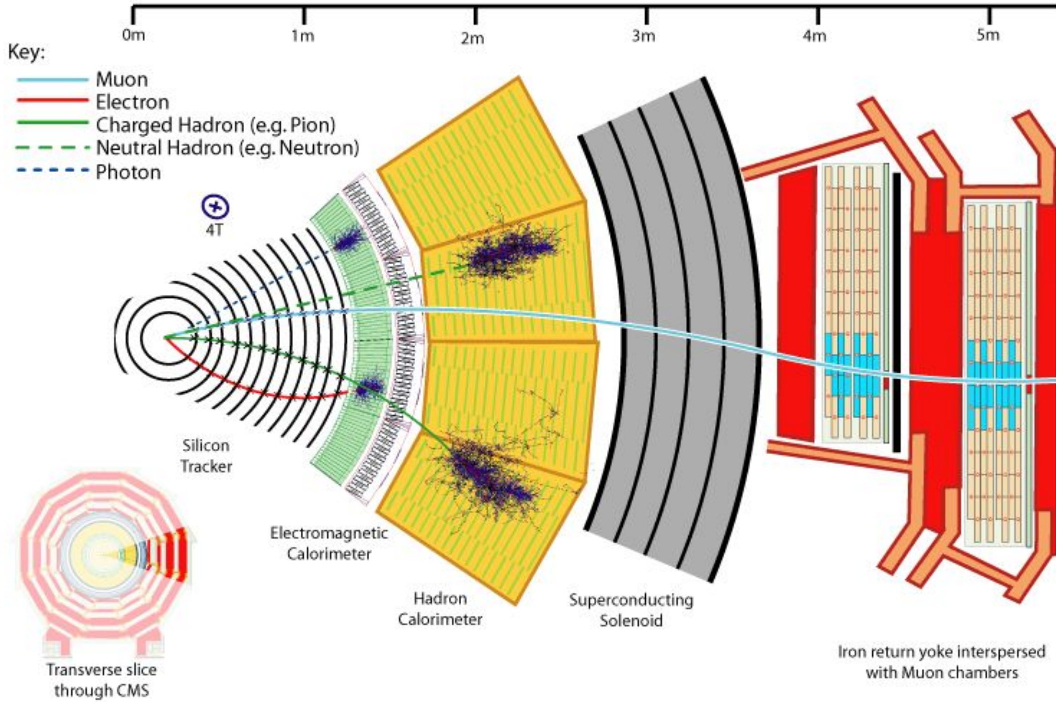
Around 76 000 **$PbWO_4$ crystals.**

Mainly used for the reconstruction of **electrons** and **photons.**

Plays crucial role for **all physics analysis**, e.g. for Higgs decay channels:

$$H \to \gamma\gamma$$

$$H \to ZZ^* \to 4\ell$$



Reconstruction in ECAL

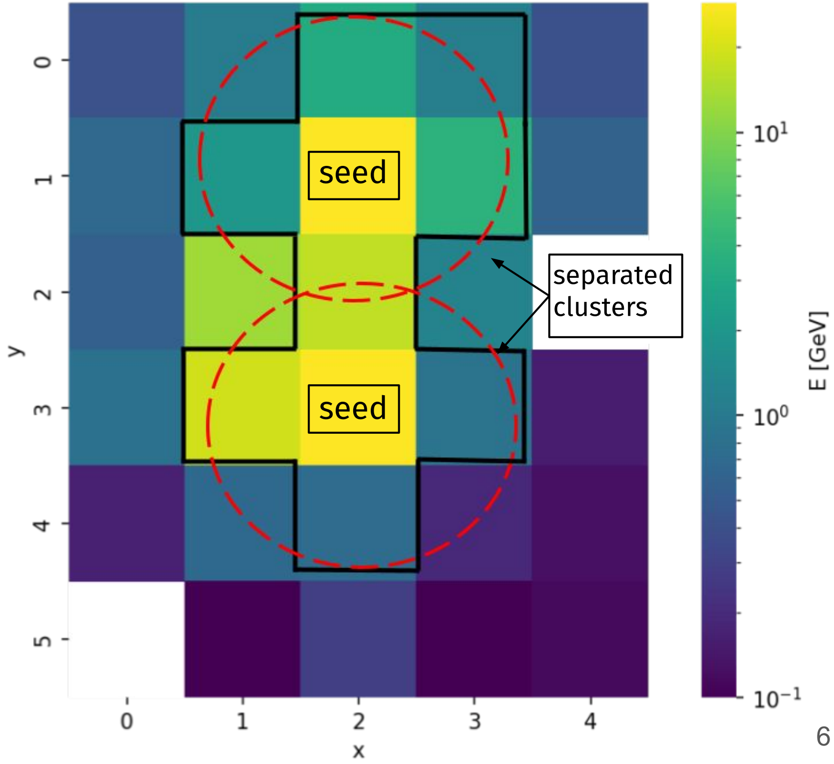
Reconstruct **position and energy** of electrons and photons from **electromagnetic showers**.

Current CMS algorithm (pfclustering):

hits — Reconstructed energy deposits left by traversing particle in crystals

clustering
- Identify local maximum (seed)
- Aggregate crystals around seed - grow a cluster
- Separate overlapping clusters using Gaussian mixture algorithm

BDT correction — Correct the predicted energy using Boosted Decision Tree



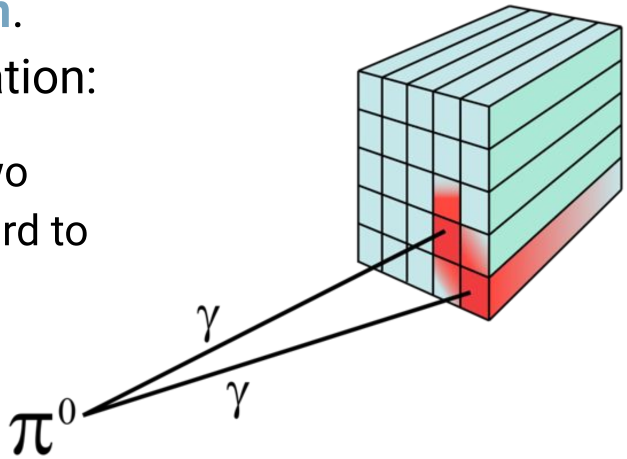
Motivation

Creating a novel **ML-based algorithm** to improve ECAL reconstruction.

Main objectives:

➔ Improving **energy** and **coordinates resolution**.
➔ Improving **photon vs. neutral pion** discrimination:

Photons coming from neutral pion decay create two overlapping clusters in the calorimeter, which is hard to discriminate from a single photon's signature.



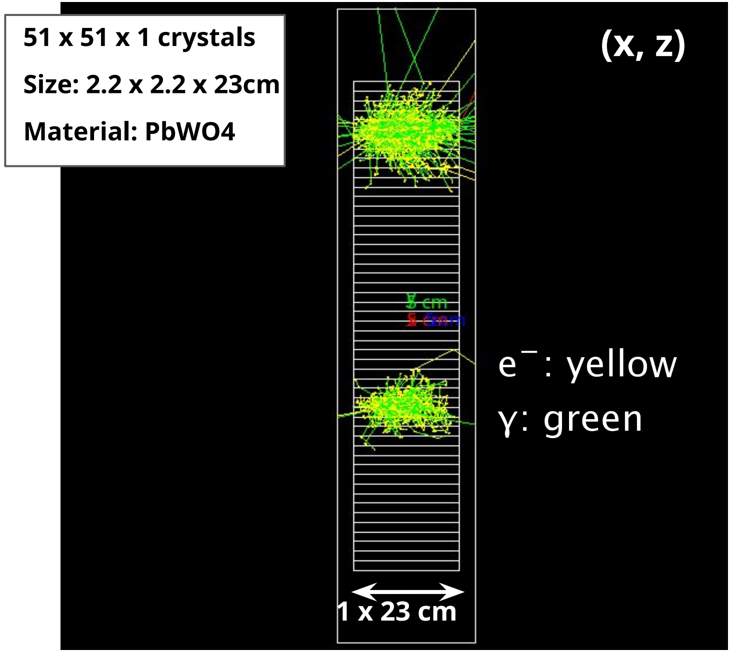
Source: science 2.0

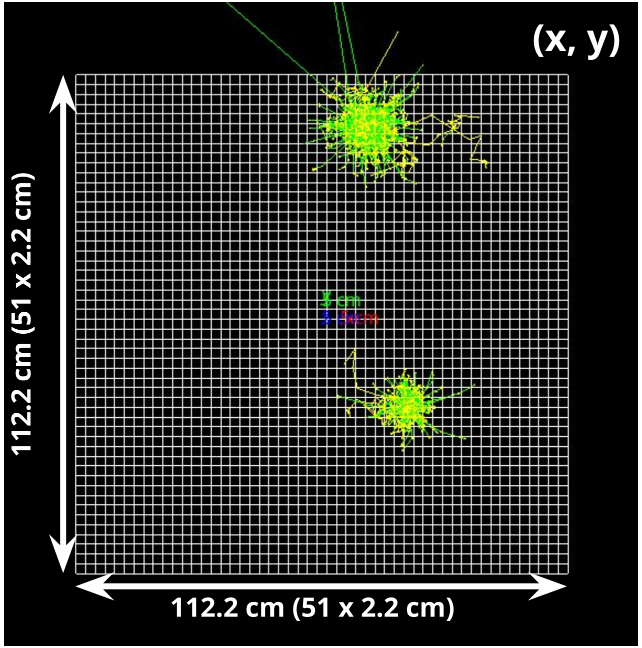
Simulation

Detector simulation

Simplified calorimeter simulated in Geant4 to test the performance of the algorithms.

Crystal parameters same as in ECAL (but not tilted).

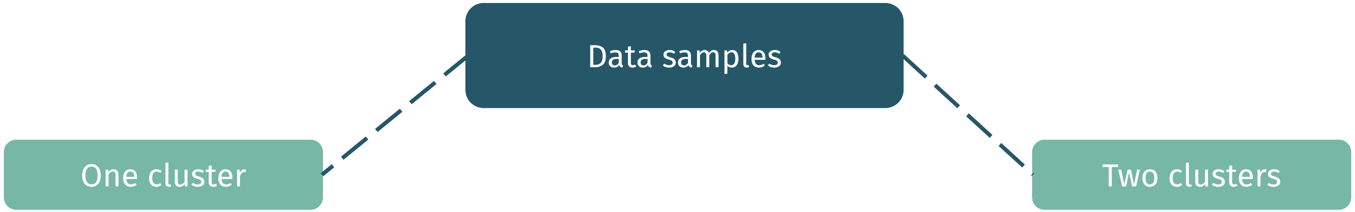


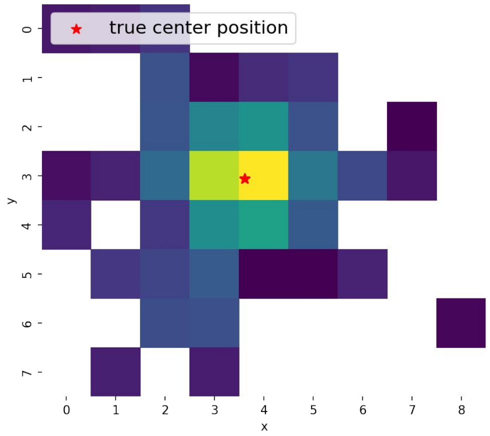


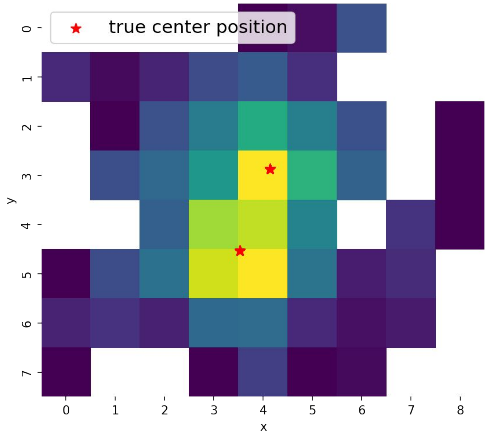
Dataset creation

Photons with [1, 100] GeV energy are used, directed perpendicularly to the calorimeter.

Per crystal noise is added (σ = 167 MeV) and cut is set on 50 MeV.







In each sample the distance between two clusters < 3 crystals.

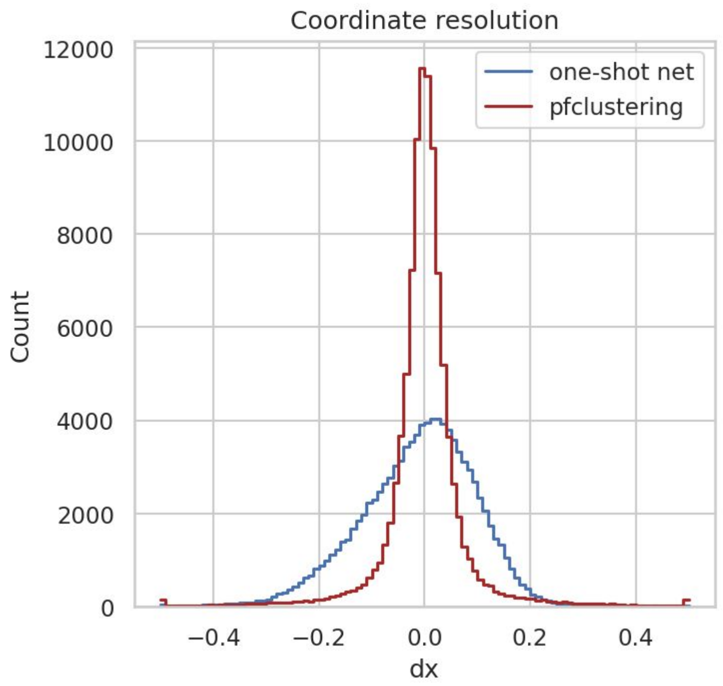
Networks

One-shot network

Energy deposits in crystals can be represented as **pixel intensities of an image** → allows to use **Convolutional Neural Network** (CNN).

First attempt:

➔ CNN applied on **full window** (dim: 51 x 51).
➔ Predicting **position of particles** (n x 2), where n - number of particles per window.



Results:

➔ The network is able to make prediction but **resolution is always worse** than for pfclustering.

Conclusion:

➔ One network is **not able to predict** number of particles and their position **simultaneously**.
➔ Not scalable for the full ECAL window (360 x 170).

Double-step network: convolutions

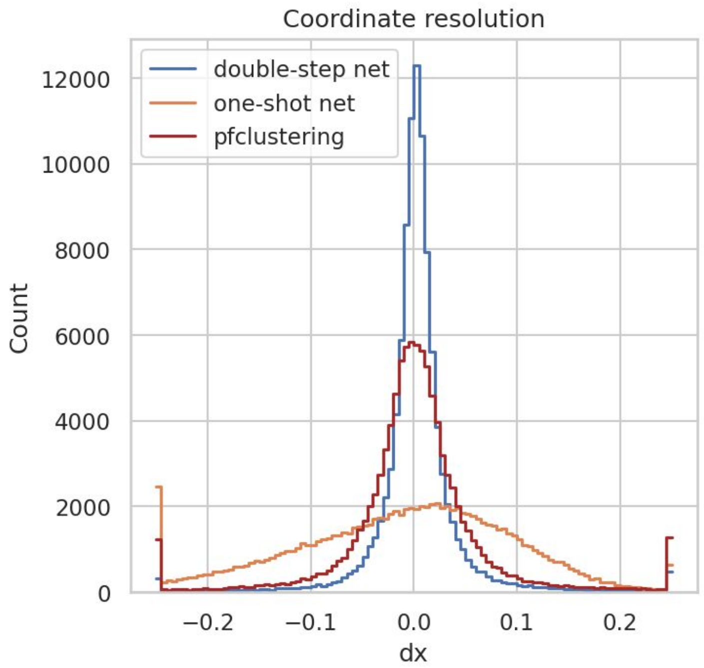
Next step: separate the task into two CNN networks (with similar architecture).

Input — 7 x 7 windows centered on all the crystals with > 0.5 GeV deposit.

1st network: seed finder — For each window predicts a probability to be a seed (or a real cluster).

Only windows with > 0.5 probability

2nd network: center finder — Predicts precise position and energy for every seed that passed the threshold.

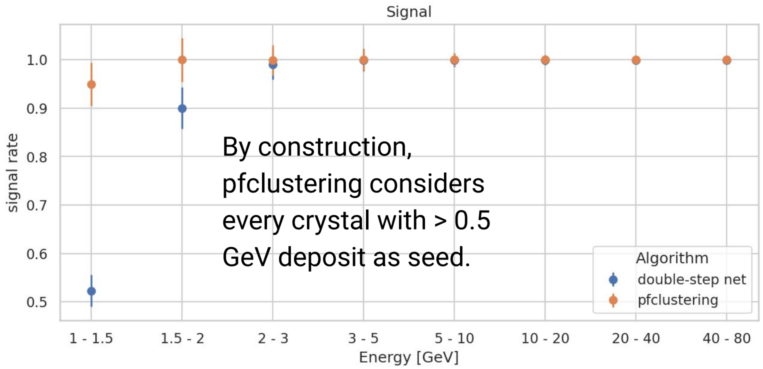


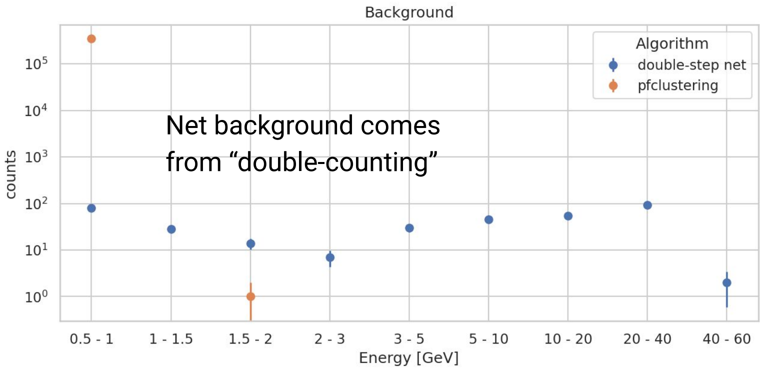
Significantly improved resolution both for position and energy reconstruction!

Double-step network: convolutions

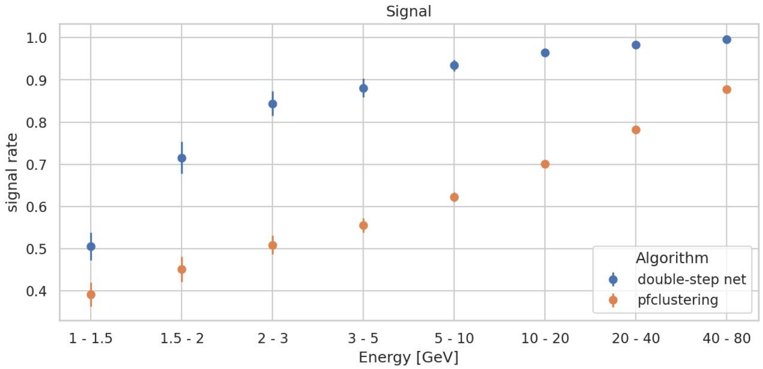
Signal - the ratio of correctly predicted clusters to the full number of clusters. Background - number of events misidentified as clusters.

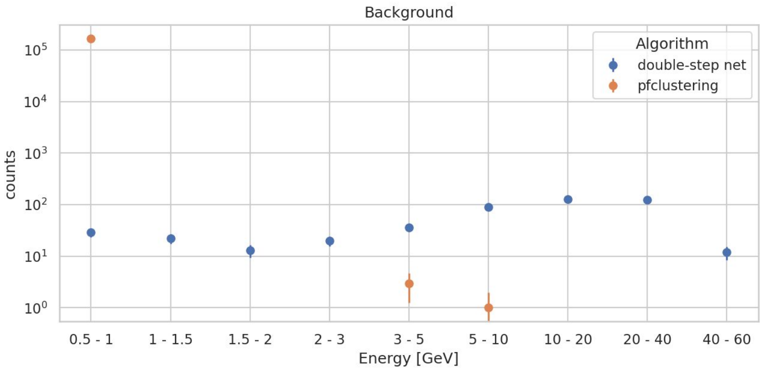
One particle





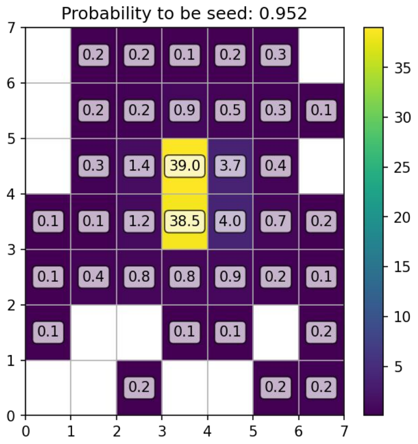
Two particles

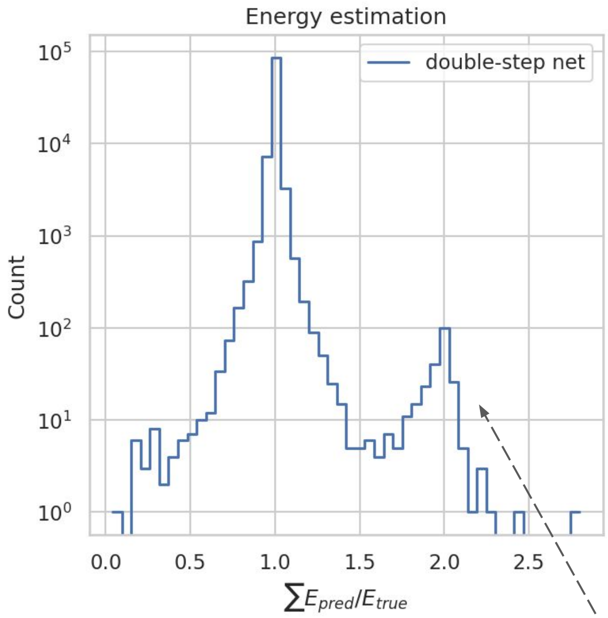




“Double-counting” problem

Input windows do not communicate in the network $\rightarrow$ problem appears when particle position is close to the border:





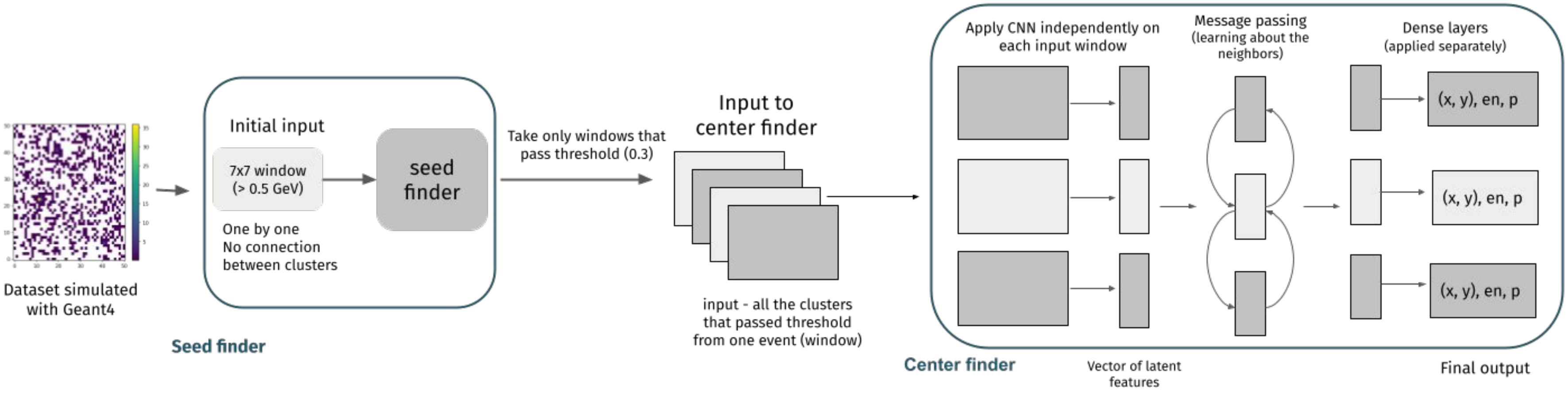
Almost identical predicted positions (12.68, 34.98) and (12.67, 35.00) as well as energies.

Creates a large energy overestimation.

Double-step network: convolutions + graphs

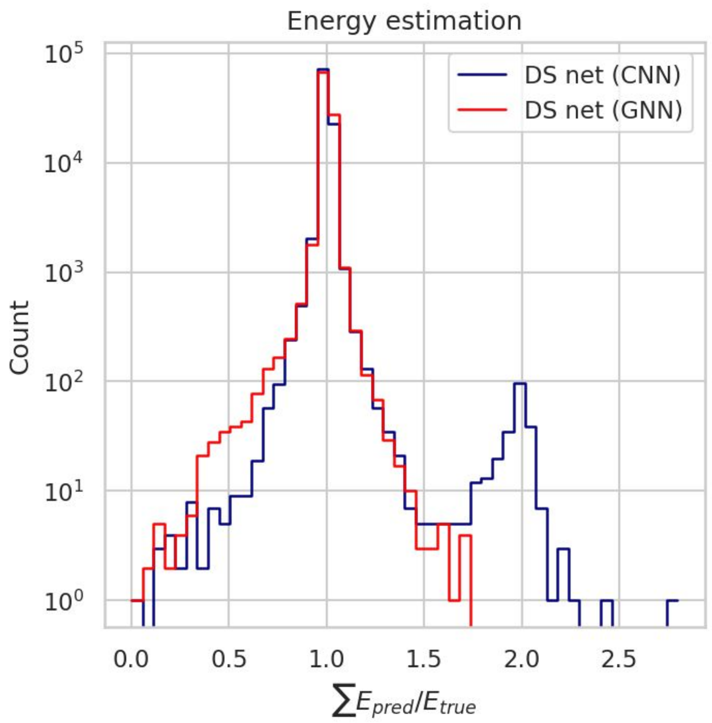
Solution: add communication between input windows.

Using **Graph Neural Networks** (GNN) and **Message-Passing** (MP) each window can **learn about its neighbors**.

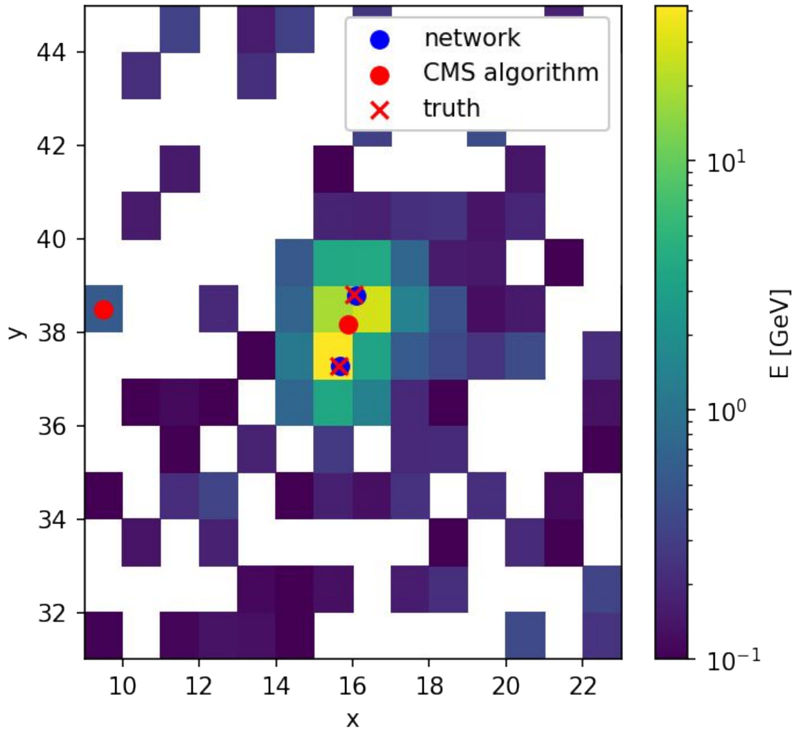


Center finder predictions are **precise coordinates, energy and corrected probability** to be a real cluster after adding MP.

Final* results



With MP overestimation of energy is significantly reduced – “double-counting” solved.

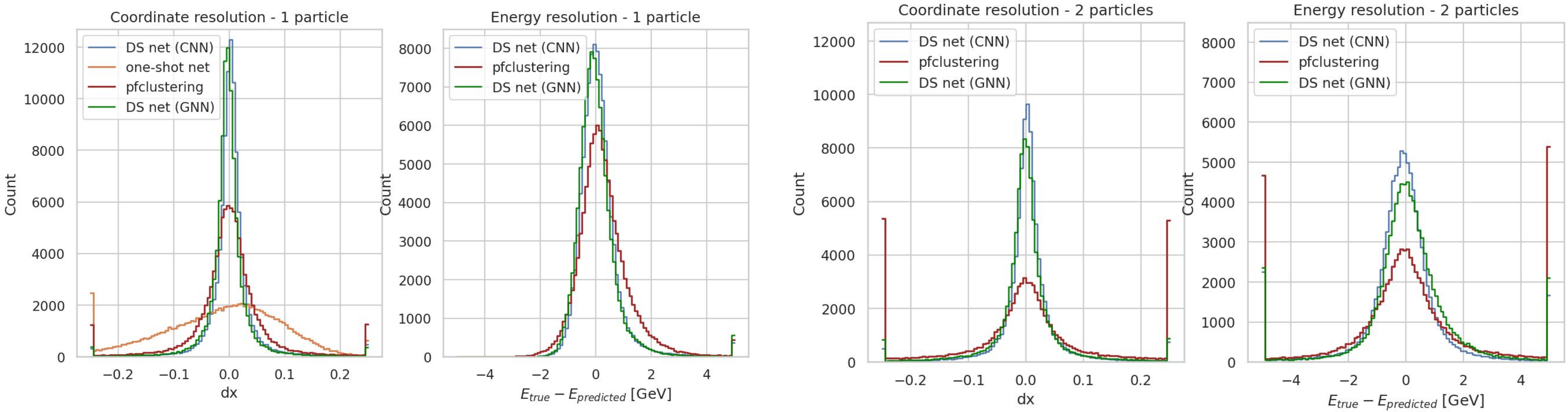


Event example where network correctly identifies two clusters while pfclustering predicts only one.

Final* results

* – convergence on the exact architecture is ongoing.

Both double-step (DS) networks perform better than pfclustering.



Significant improvement in resolution: 0.05 vs. 0.08 ECAL crystals and 0.54 vs. 0.71 GeV !

Outlook

Short-term objectives:

➔ Finalizing the network architecture.
➔ Testing performance on “more than two” clusters per window.
➔ Publishing a paper on the achieved results.

Long-term objectives:

➔ Implementing the ML algorithm in CMS software and estimating the effect on the final physics analyses.

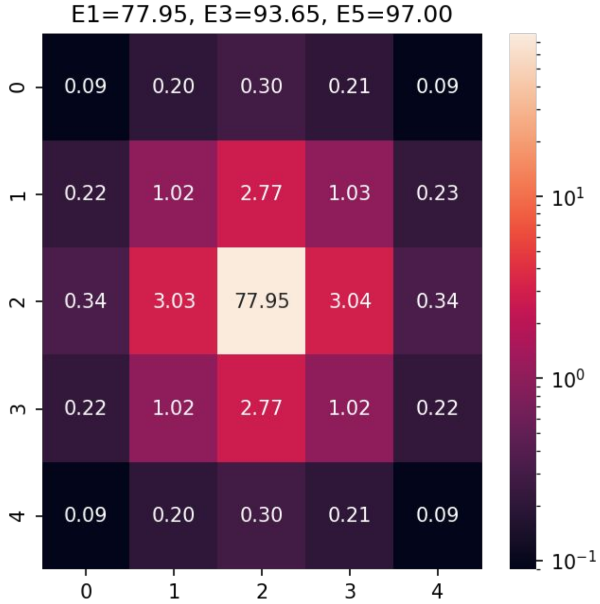
Source: explain xkcd

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No *800945* — NUMERICS — H2020-MSCA-COFUND-2017

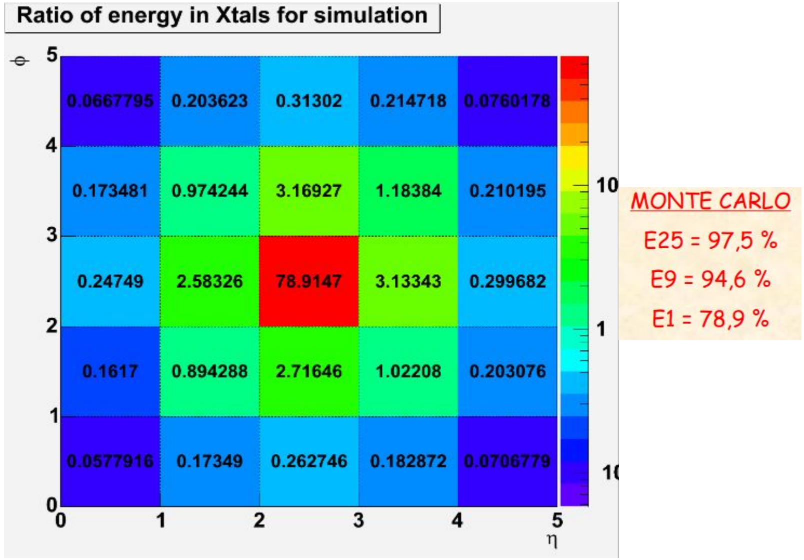
Backup

Energy deposit profile

- To **validate** the calorimeter simulation the **energy deposit profile** was plotted.
- Using the data from **1 000 electrons, all at 100 GeV** shooting at the **crystal center of the middle crystal.**



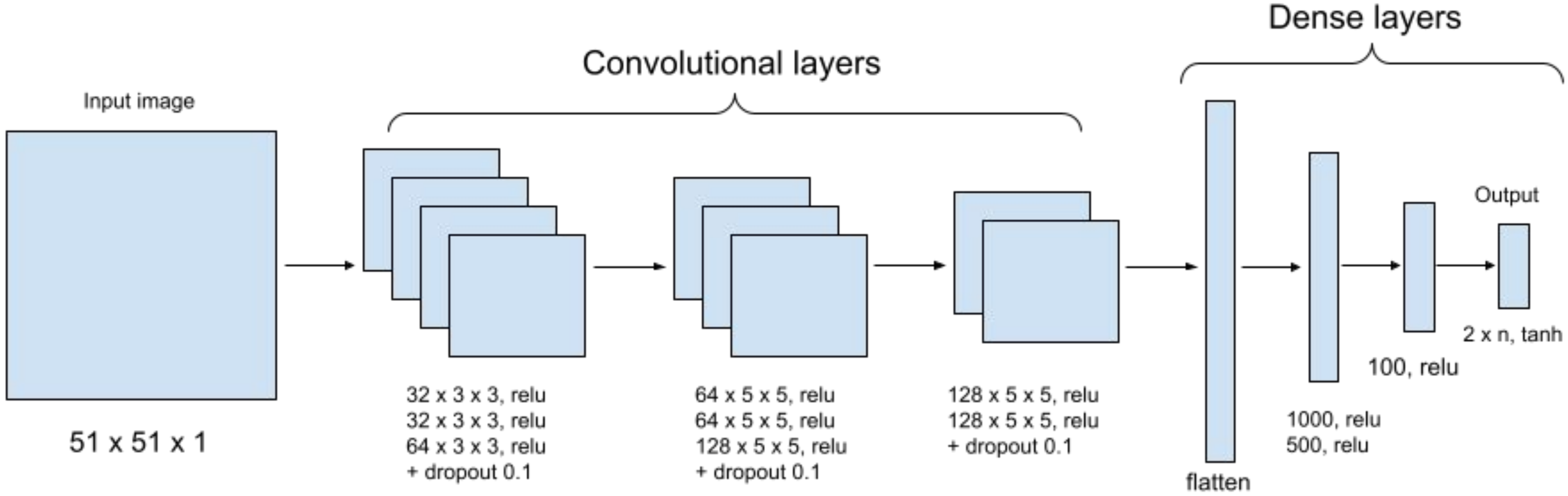
Energy deposits from the simplified detector.



Energy deposits from Geant4 simulation of ECAL.
http://geant4.in2p3.fr/2005/Workshop/UserSession/P.Mine.pdf

The results are very similar => the simulation can be used as a **proxy for CMS ECAL**.

One-shot network architecture



Hyperparameters:

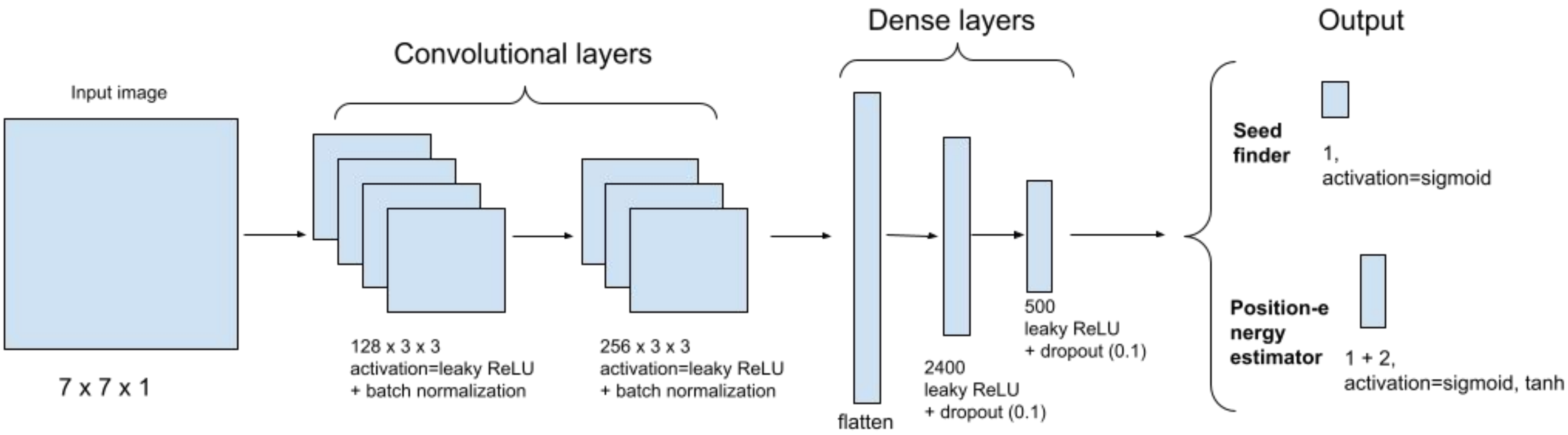
learning rate = 0.001

batch size = 64

epochs ~ 500

Loss function: Mean Absolute Error

Double-step network architecture



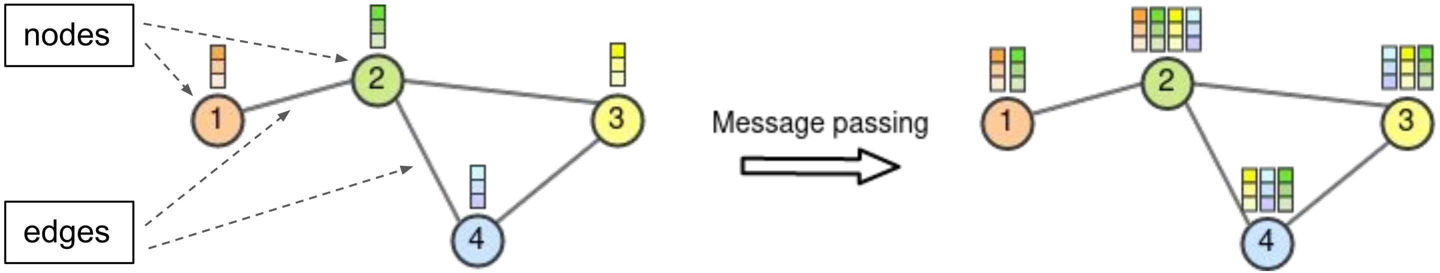
Hyperparameters:

learning rate = 0.0001
batch size = 64
epochs ~ 500

Loss function: Binary Crossentropy (for seed finder) or Mean Absolute Error (for position-energy estimator).

Graph Neural Networks

- ➢ Type of neural network that can operate on and analyze **graph structures**.
- ➢ Unlike other types of networks GNN can be easily applied on sparse data, doesn't require padding.
- ➢ A graph consists of **nodes** (contain features of the object) and **edges** (reflect the relationship between the nodes).
- ➢ In GNNs the information can be shared between the neighbors:
  - ○ The vector features of each node are transformed into "messages" (e.g. using dense layers) that are sent to the neighbors (message-passing).
  - ○ In this way, **each node learns information about its neighbors and itself**. The process is carried out in parallel and repeated several times.



https://uvadlc-notebooks.readthedocs.io/en/latest/tutorial_notebooks/tutorial7/GNN_overview.html