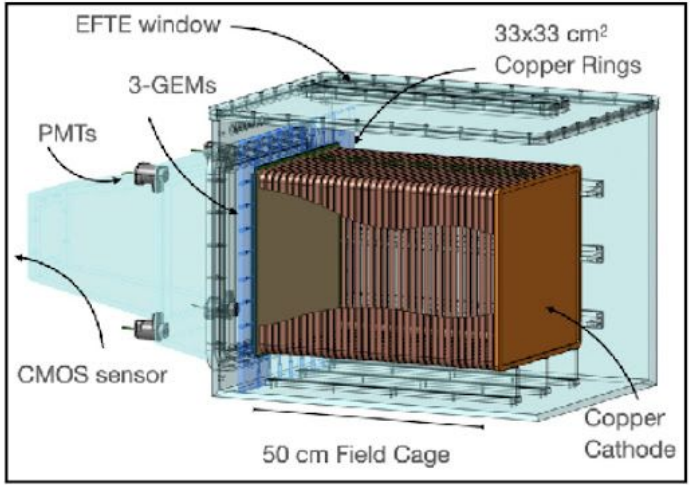
Position Reconstruction from CYGNO LIME PMTs

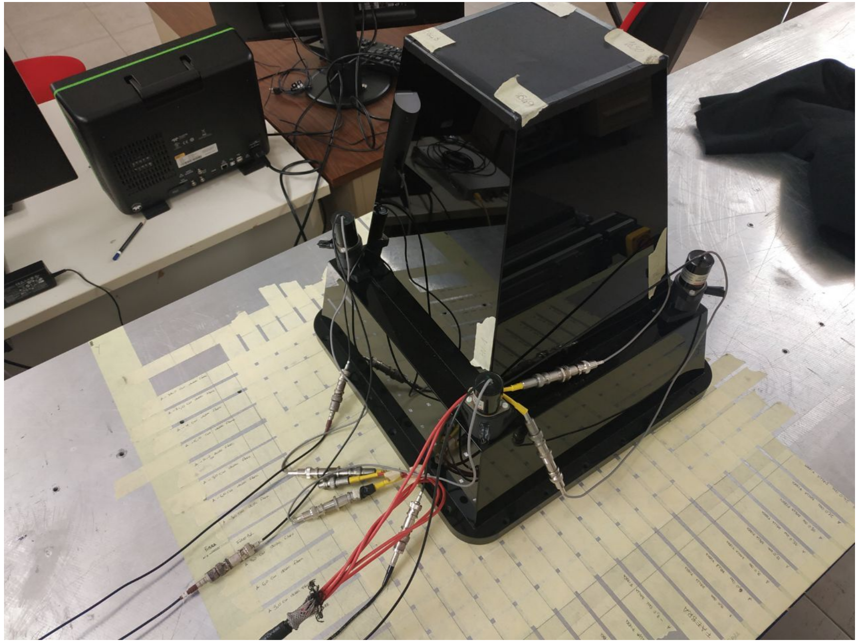
Luca Zappaterra

Overview and main goals

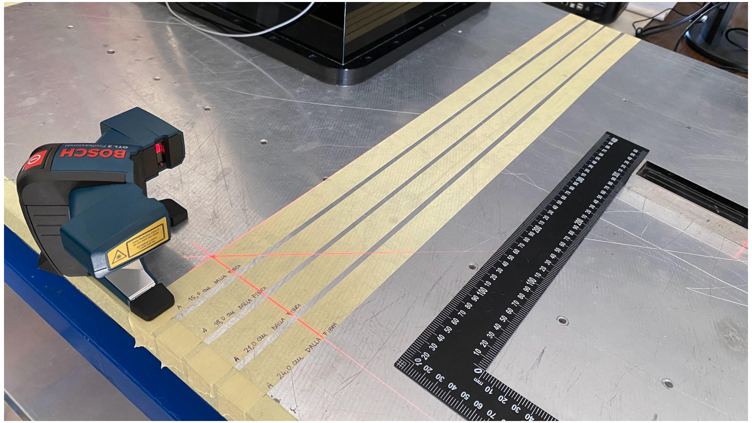
- ➢ The INFN **CYGNO experiment** aims to unveil the Dark Matter mystery by means of **Directional DM search**, pointing its detectors towards the Cygnus constellation.
- ➢ One of these detectors is called LIME1, a directional TPC with optical readout (**sCMOS**) and fast light sensors (**PMTs**).
- ➢ The joint work of both sCMOS camera and PMTs <u>**should**</u> allow the 3D reconstruction of particles tracks inside the detector.
- ➢ This project aims to facilitate this operation, with the development of a **Neural Network for 2D position reconstruction** (the particle entrance position) via time-series analysis of PMT waveforms.



With great datasets comes great sensibility



- A light pulse generator (**LPG**) have been used to simulate a particle event, while the PMT housing structure (that we will call the **cone**) has been slid over a labeled tape grid.
  - *Great thanks to my lab mates Daniele, Giulia and Pierfrancesco, which helped me in this operation!*
- A total of **82 points collected**, spaced by 3 cm in both x and y directions, 1000 events each.
  - 64% used as training set, 16% as validation set and 20% as test set for the Neural Network

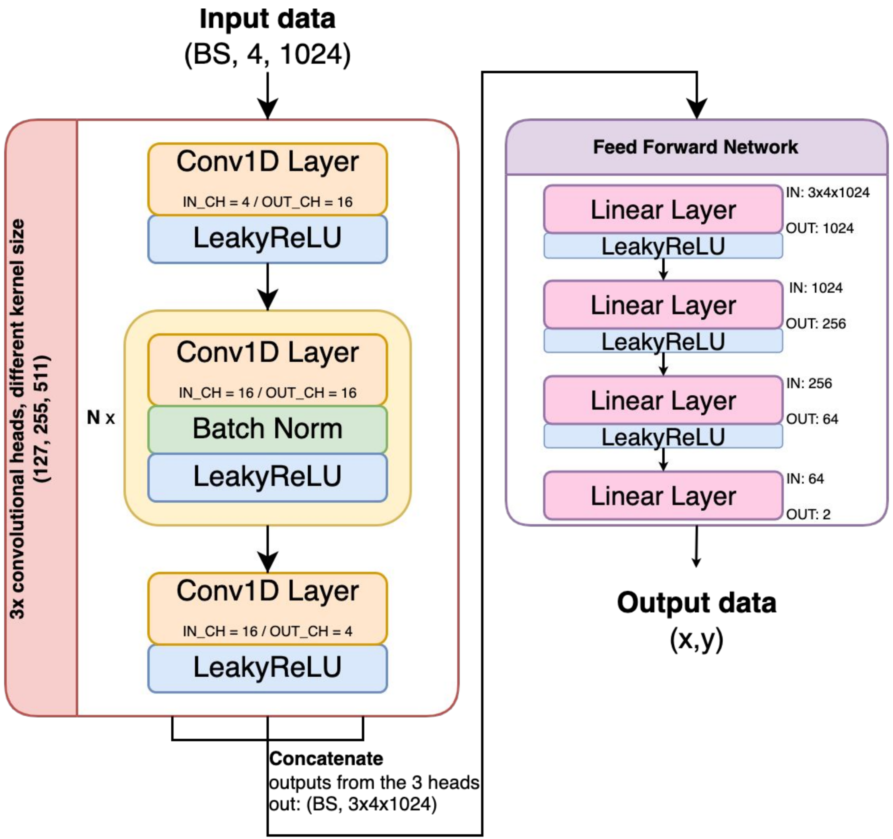


- **Above**, the final DAQ setup.
- **On the left**, the grid making-of.

Neural Network Architectures

Two models:

- **MLP** (four fully-connected layers)
  - Really simple
  - Easy and fast to train
  - Good Results
- **MHCNN** [Multi-Head Convolutional + MLP] (scheme on the right with N=2)
  - More complex and slower to train
  - Seems to be able to extract relevant information from different parts of the input
  - Results are a little better than MLP and the network is able to predict every position with 100% accuracy when approximating to integer values



Main Results

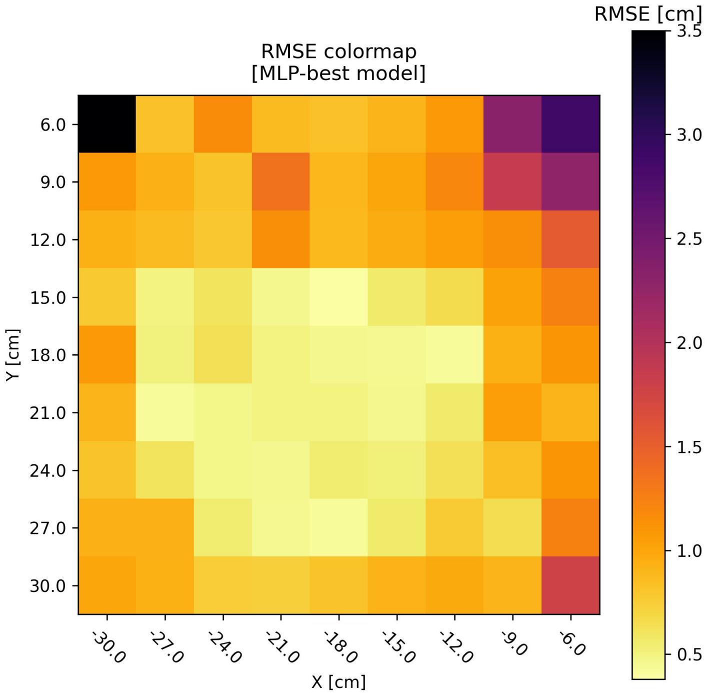
Both networks were trained with *MSE/HuberLoss* as loss function and *Adam* as optimizer.

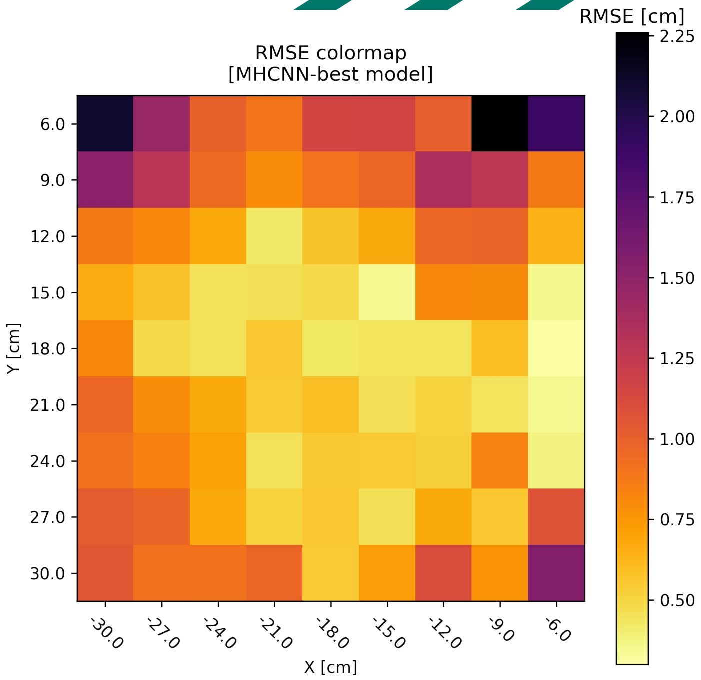
A **substantial improvement** in performance was obtained **normalizing each dataset** to the maximum value in that dataset (in order to maintain the correlation between events related to a certain position).

- **MLP** (total epochs = 50) - approx. 1 Million parameters
  - best model @ 47 epochs, HuberLoss, L.R.=10^{-3} (approximately **2 minutes** of training using the **full dataset**).
- **MHCNN** (total epochs = 15) - approx. 13 Million parameters
  - best model @ 15 epochs, MSELoss, scheduler to scale the L.R. from L.R.=10^{-3} by 1/10 after 10 epochs (approximately **30 minutes** of training using only **half of the dataset**).

➔ Again, results are similar, but the MHCNN architecture offers a much wider range of improvement by combining *refined hyperparameters tuning*, *higher training time* and *larger datasets*.

RMSE colormaps in the cone area are shown for comparison [**mind the colorbars!**]





Tests results examples

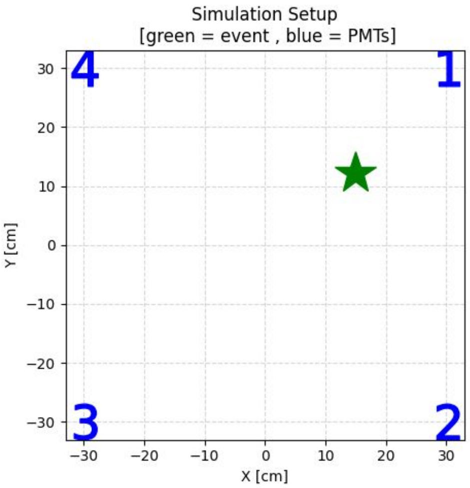
```
Best model for MLP was saved at 47 epochs

Prediction: [[-20.62  14.97]
 [-27.03  23.75]
 [-26.95  30.05]
 [-24.05   6.07]
 [ -9.08  12.01]
 [ -8.91  26.92]]
Target: [[-21.  15.]
 [-27.  24.]
 [-27.  30.]
 [-24.   6.]
 [ -9.  12.]
 [ -9.  27.]]
Loss: 0.06471997499465942
rmse: 0.25440120697021484
Test accuracy: 93.36 %
```

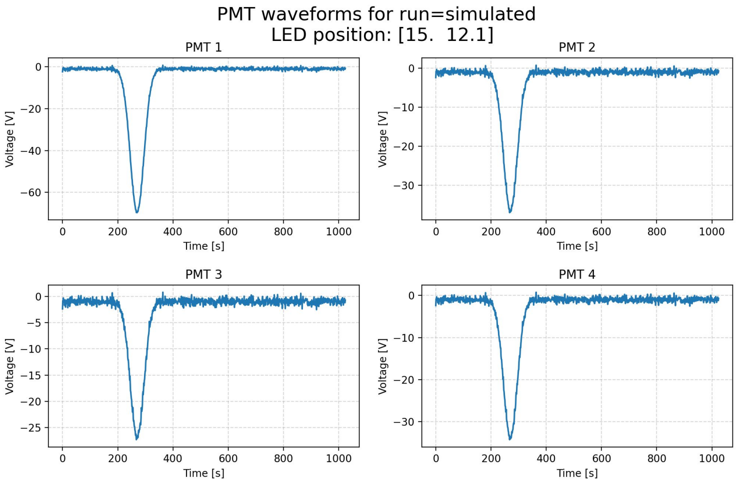
```
Best model for MHCNN was saved at 15 epochs

Prediction: [[-26.96  27.07]
 [-27.05  15.07]
 [-17.97  17.89]
 [ -9.02  26.87]
 [ -8.94  20.99]
 [ -5.98  18.03]]
Target: [[-27.  27.]
 [-27.  15.]
 [-18.  18.]
 [ -9.  27.]
 [ -9.  21.]
 [ -6.  18.]]
Loss: 0.01172076165676117
rmse: 0.10826246440410614
Test accuracy: 100.00 %
```

Toy model simulation

- A toy model has been constructed to **simulate the passage of a particle at a random position and the respective waveforms seen by the PMTs**. The latter are modeled as gaussians (eventually with gaussian additive noise), scaled as 1/r , where r is the relative distance between a PMT and the “particle”, to emulate the signal attenuation with increasing distance from the event point.

- On the right an example of a simulated event and related waveforms is shown
- From recent results, obtained from PMT equalization analysis carried out during lab activities, **it has been possible to improve the simulated PMT response.**

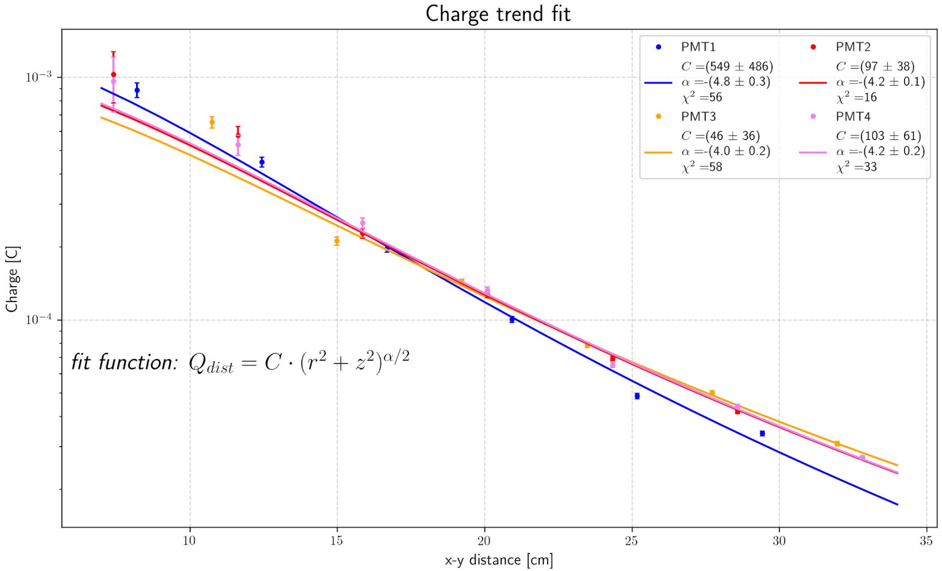




Toy model simulation improvements - hints

- Integrating the waveforms of the PMTs along the diagonals of the **real** scanned area, it was possible to extrapolate the light isotropy and quantitatively determine the charge trend by means of 4 fits; as shown in the figure below, **the four trends are compatible** and show a clear $1/r^4$ scaling.

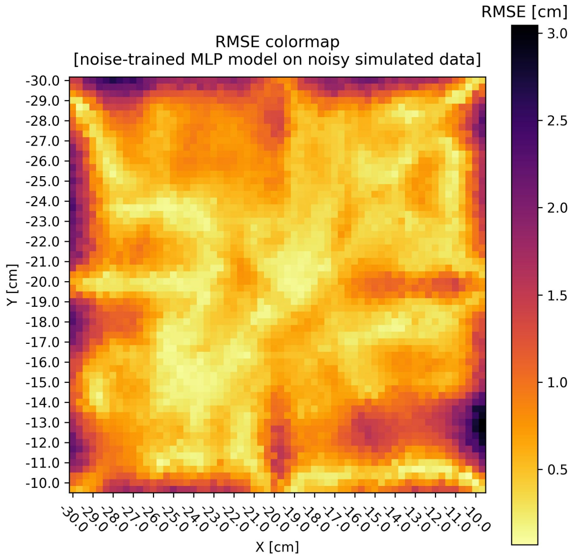
- By means of a **global fit** (justified by the previous analysis), the toy model was improved by imposing the PMT height with respect to the grid (***z=10 cm***) and the correct scaling factor (**$1/r^4$**).

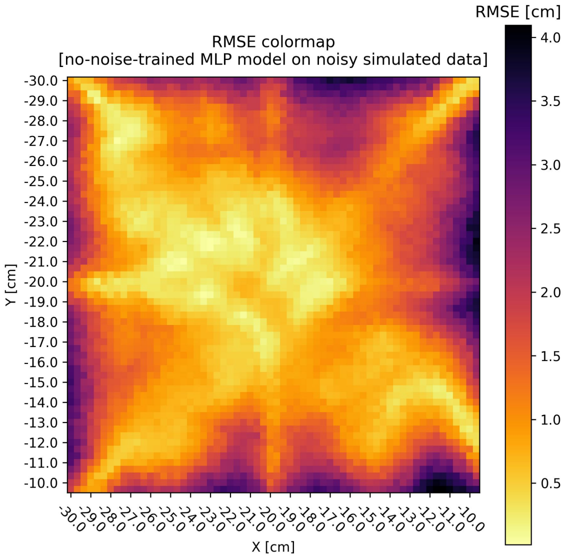


Toy model test results [MLP]

- The Neural Networks have been trained over the simulated dataset (both noisy and clean) and RMSE maps have been computed testing the trained models on **noisy** data, to highlight eventual discrepancies in their behaviour.
- These maps are almost continuous with respect to the previous ones, since **10^5 points were generated** inside the square, assimilating the process to a MC simulation.
- From the colormaps it is clear that the model trained on clean data is way less accurate than the other, especially on the borders.

  Nevertheless, around the diagonals and the center the two models are almost equivalent.

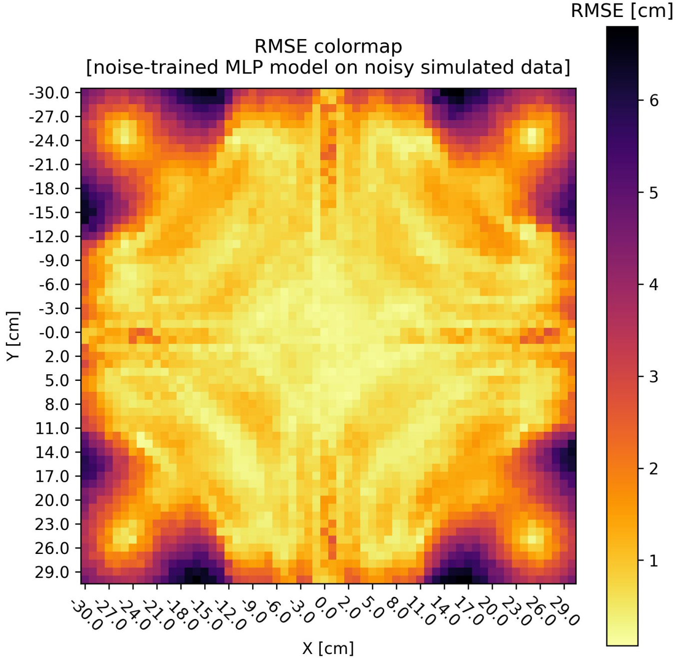




Improved Toy model test results [MLP] 2/2

- When changing the scaling of the simulated waveforms from 1/r to $1/r^4$, the results are way different, and show a very peculiar colormap.

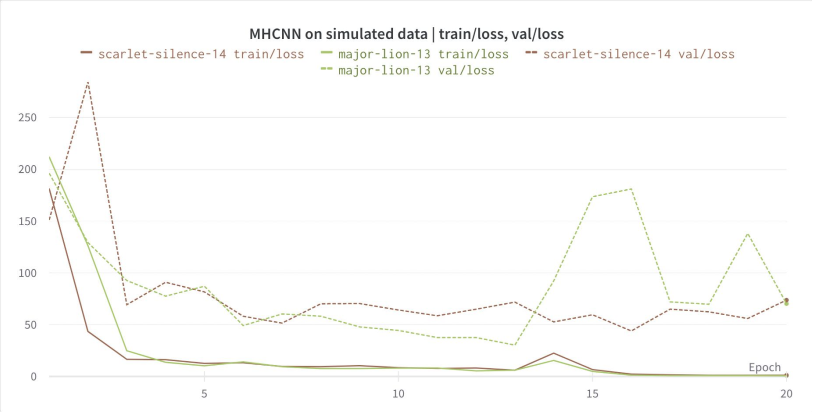
- The PMT areas are visible and the RMSE in these positions is almost zero, due to the fact that the waveforms must have a really big peak for a single PMT, clearly recognized by the network.
- Near these areas, there are two equivalent positions (diagonally symmetric) where the network has the highest RMSE; this behaviour can be associated with some geometrical symmetries of the modeled waveform trend.



Toy model test results [MHCNN]

- Testing the **MHCNN architecture on simulated data** with different hyperparameters setup (included kernel sizes) resulted in an overall **overfit** of the network.
- This overfit could be related to the intrinsic artificial (Gaussian) nature of the data, which the network was able to assimilate and recognize.

- This last statement is supported from the really good results obtained by this architecture on the real data and could be a **signal of the correct ability of the network to extrapolate** from the waveform the **physical information** useful to reconstruct the position.



Name (2 visualized)	State	Runtime	epochs	events_per_iteration	hidden layer channels	hidden layers	kernel_sizes	learning_ra	loss_func	metric	optimizer	scheduler step, gamma	train/loss	val/loss
scarlet-silence-14	finished	6m 54s	20	1000	16	4	[25,99,199]	0.001	MSELoss()	L1Loss()	Adam	[15,0.1]	1.153	73.72
major-lion-13	finished	7m 32s	20	1000	32	4	[49,99,199]	0.001	MSELoss()	L1Loss()	Adam	[15,0.1]	0.6699	70.253