Neutrino Physics and Machine Learning 2024

CNN for track reconstruction and PID in the new HA-TPCs of the T2K near detector

Anaëlle Chalumeau

NPML2024 – 27/06/2024

git repo

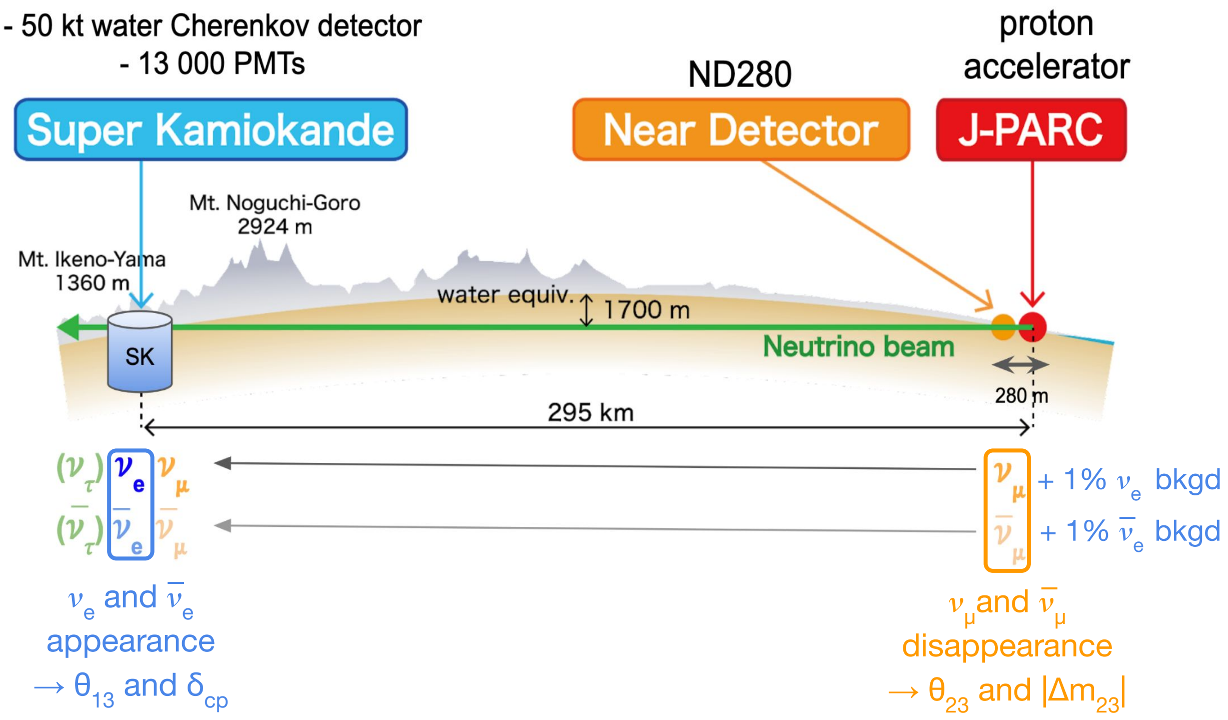
technote

Overview

- **T2K and its Near Detector ND280**
- CNN for track reconstruction (the theory)
- CNN for track reconstruction (in practice)
- Results on momentum reconstruction
- Results on Particle IDentification

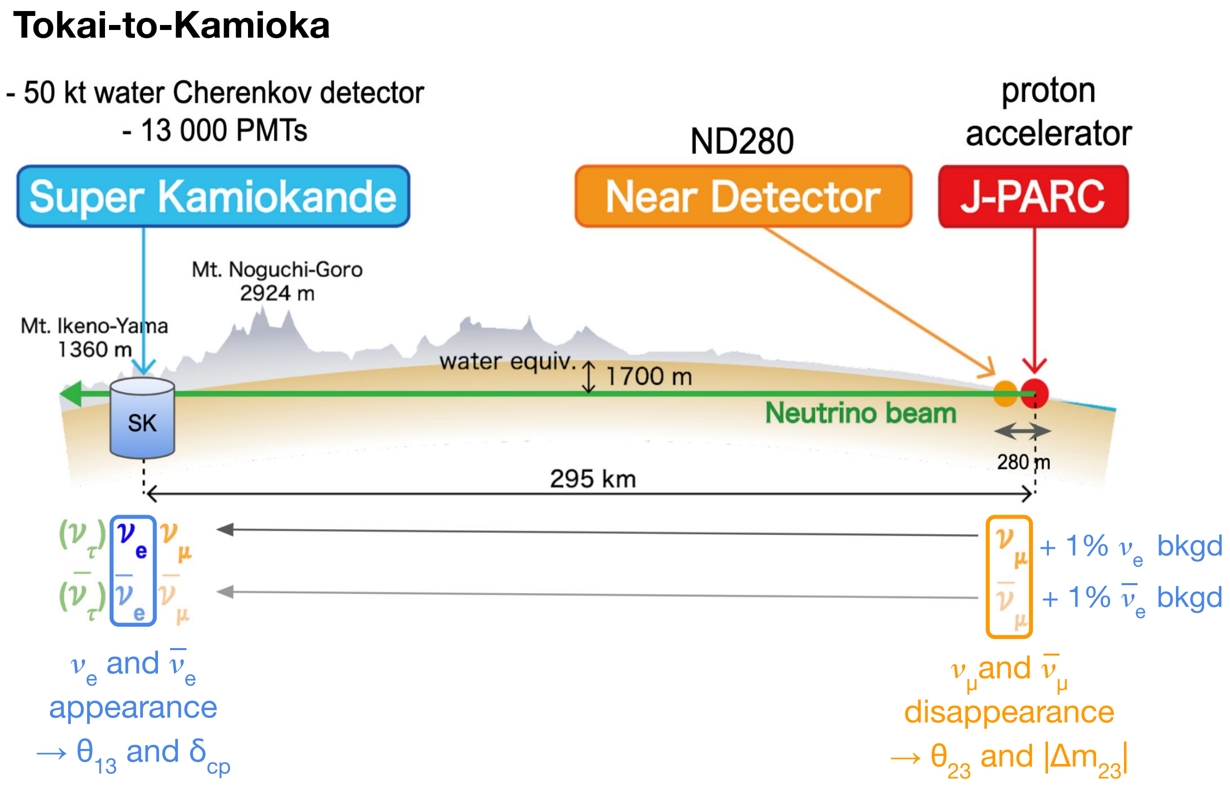
The T2K experiment and its Near Detector

Tokai-to-Kamioka



- T2K detects neutrinos at both the **Near Detector ND280** and at the **Far Detector SK** to study neutrino oscillations

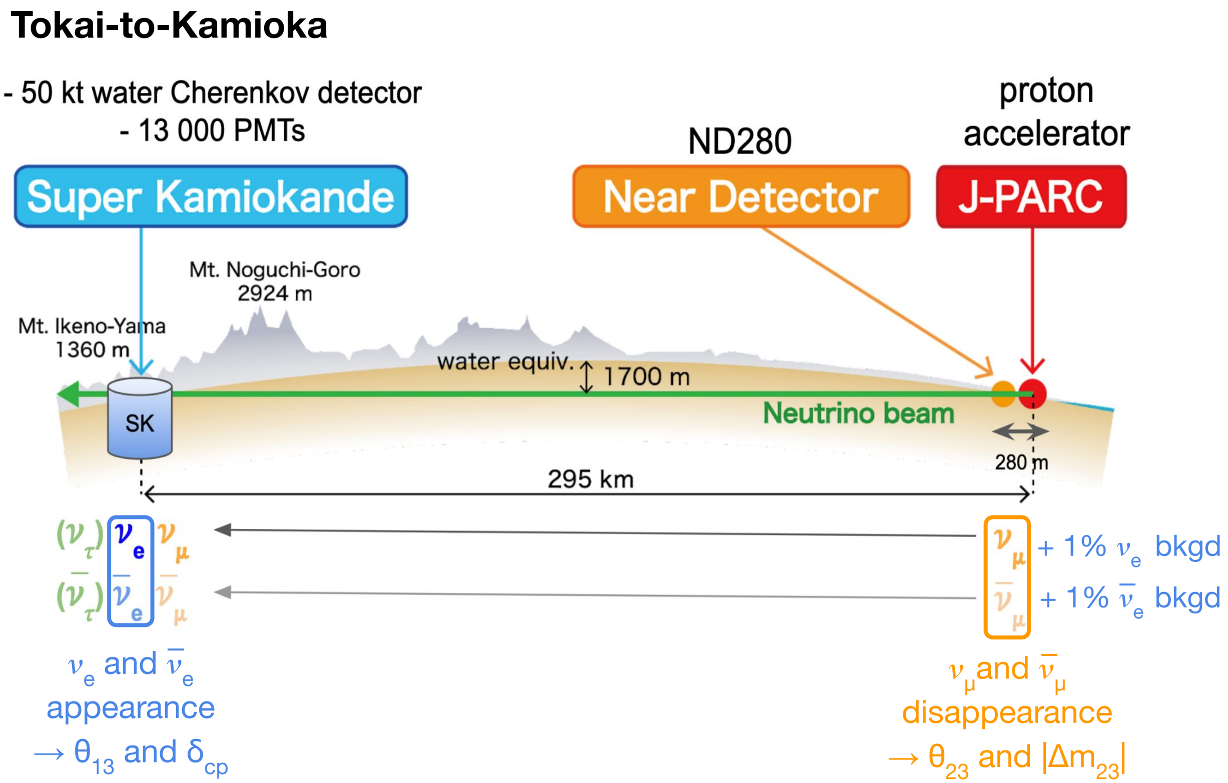
The T2K experiment and its Near Detector



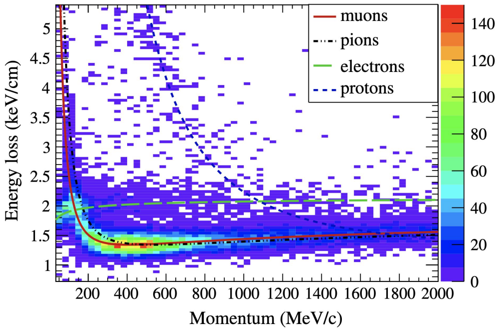
- T2K detects neutrinos at both the **Near Detector ND280** and at the **Far Detector SK** to study neutrino oscillations

- Near Detector: measurement before oscillation of the **beam spectrum** and **flavor composition**

The T2K experiment and its Near Detector

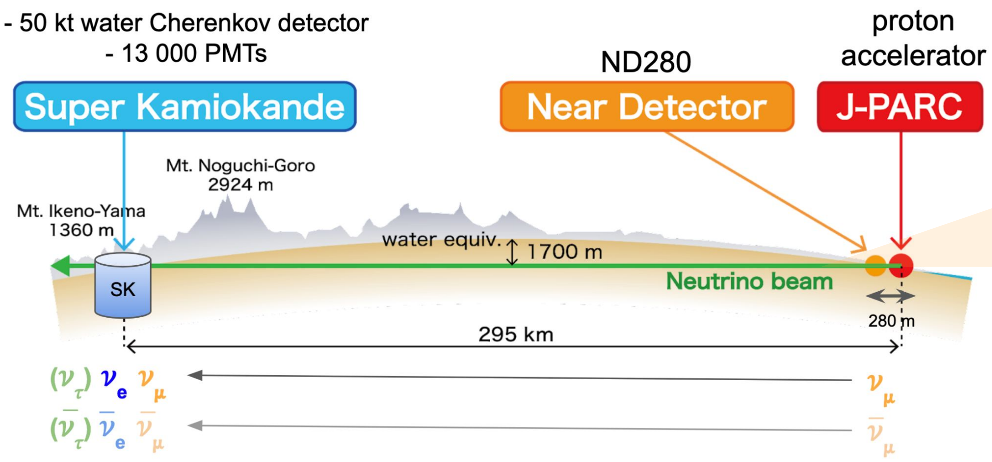


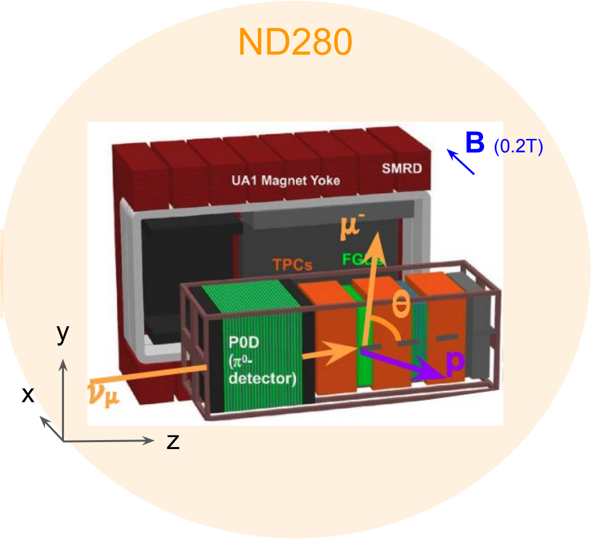
- T2K detects neutrinos at both the **Near Detector ND280** and at the **Far Detector SK** to study neutrino oscillations
- Near Detector: measurement before oscillation of the **beam spectrum** and **flavor composition**
- Need precise measurement at ND280, e.g. to **distinguish the ν_e bkdg from the ν_μ signal**



The T2K experiment and its Near Detector

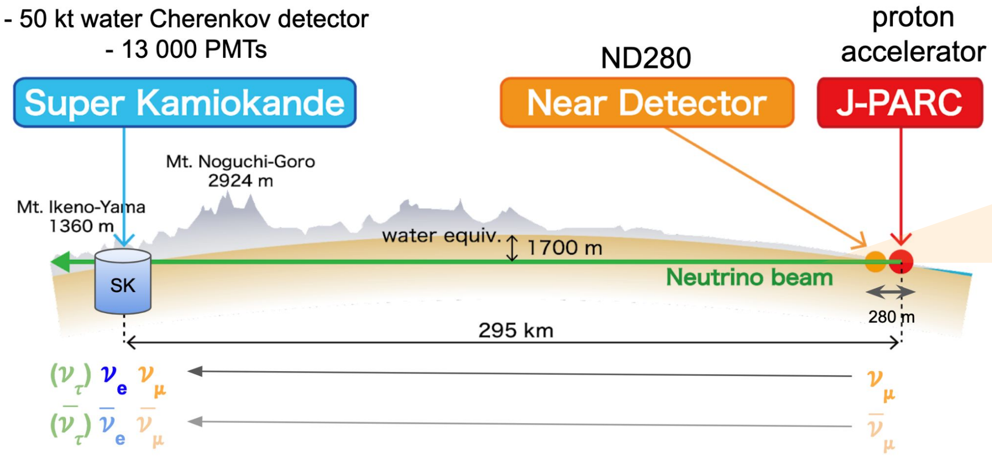
ND280 and its HA-TPCs

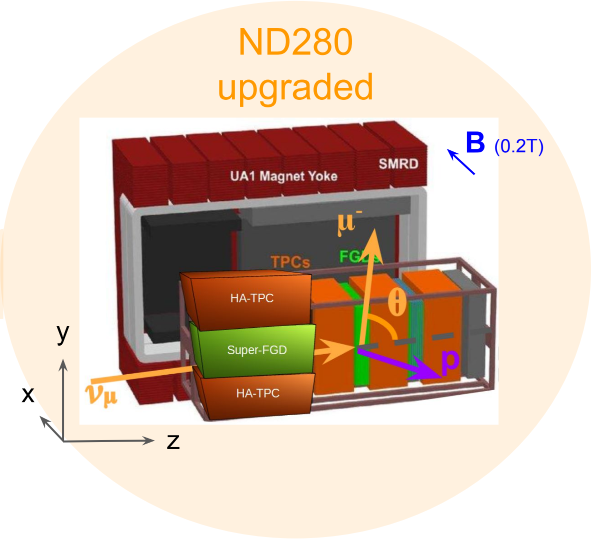




The T2K experiment and its Near Detector

ND280 and its HA-TPCs

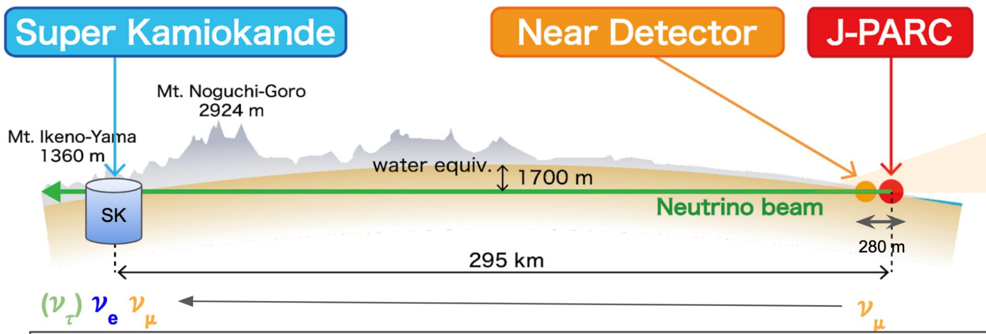


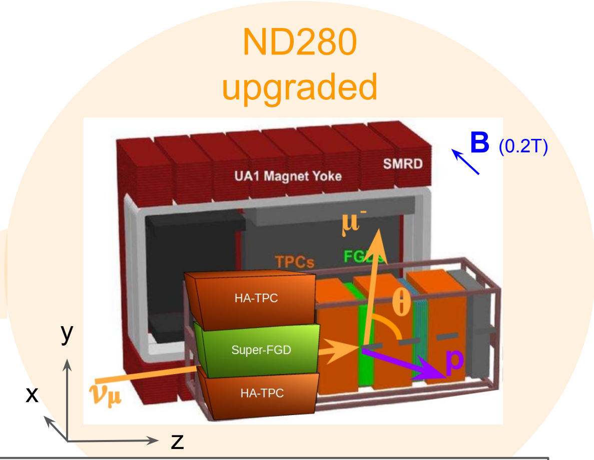


The T2K experiment and its Near Detector

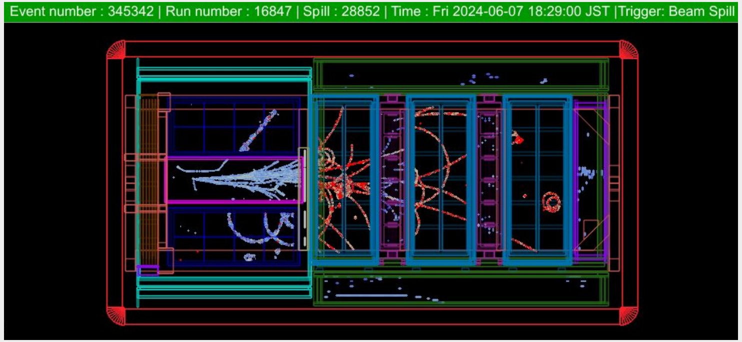
ND280 and its HA-TPCs

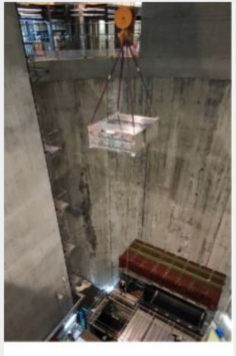
- 50 kt water Cherenkov detector
- 13 000 PMTs



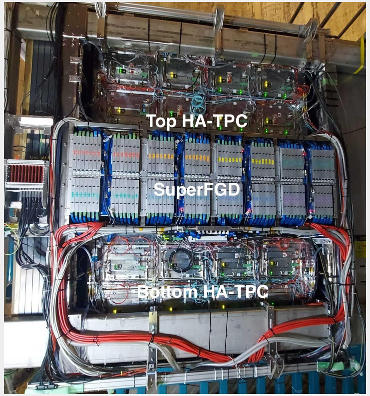


upgrade installed and taking data right now!



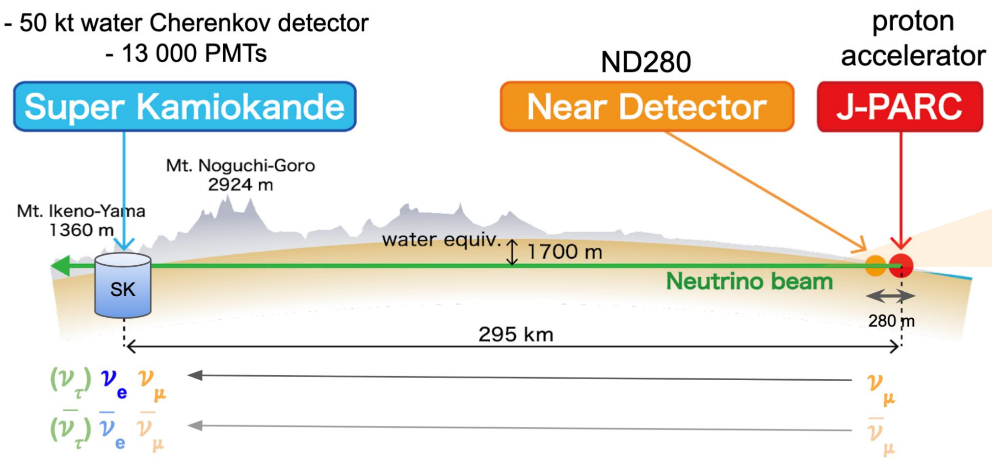


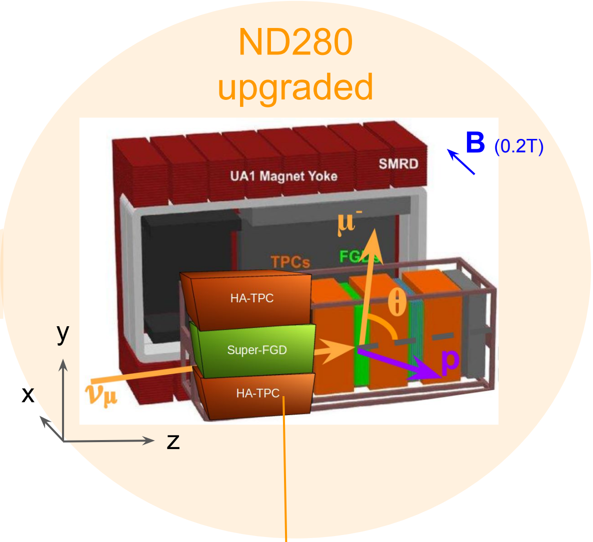
Lowering bottom HATPC
2023.9.8

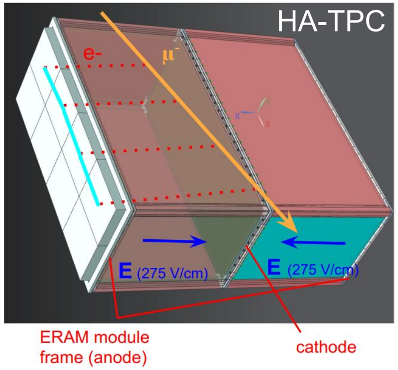


The T2K experiment and its Near Detector

ND280 and its HA-TPCs

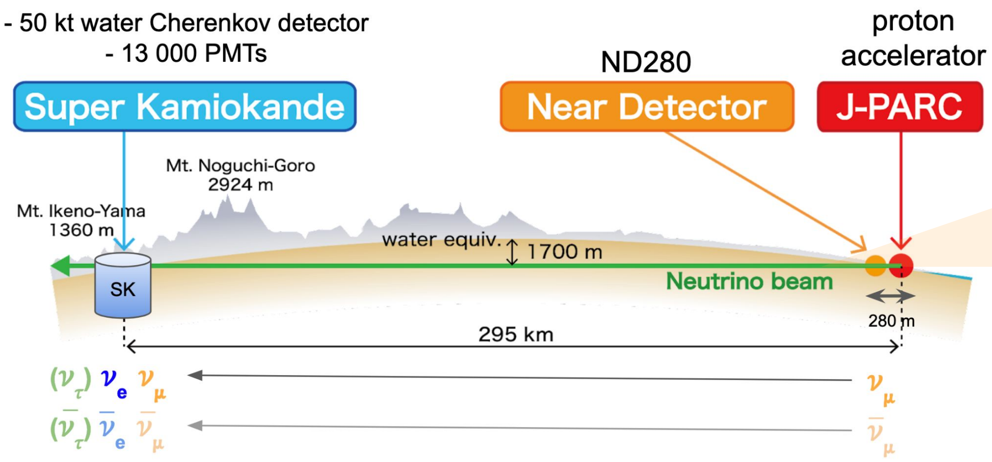


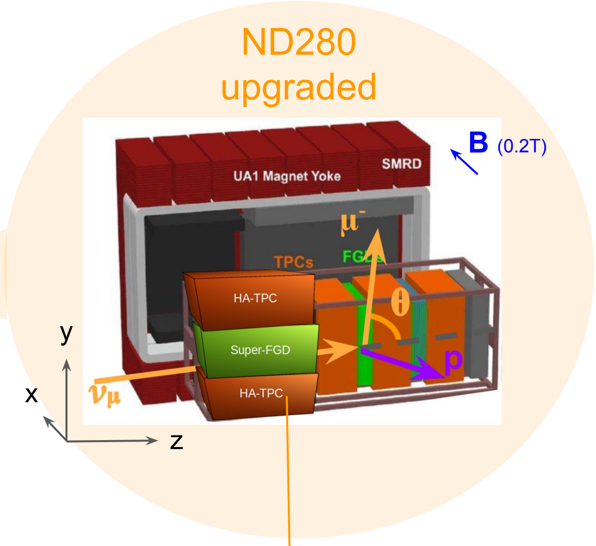




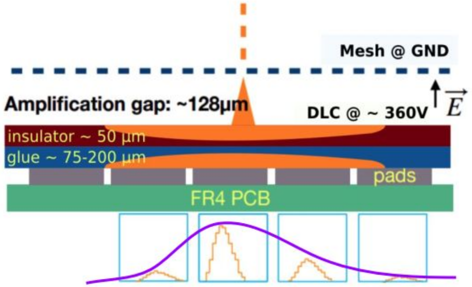
The T2K experiment and its Near Detector

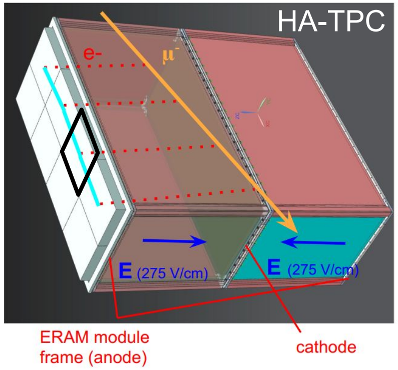
ND280 and its HA-TPCs





Encapsulated Resistive Anode MicroMegas (ERAM)

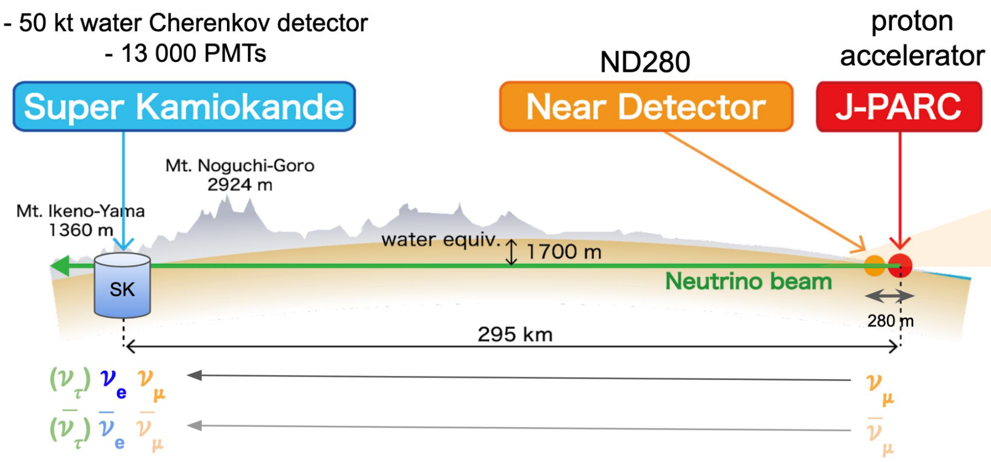


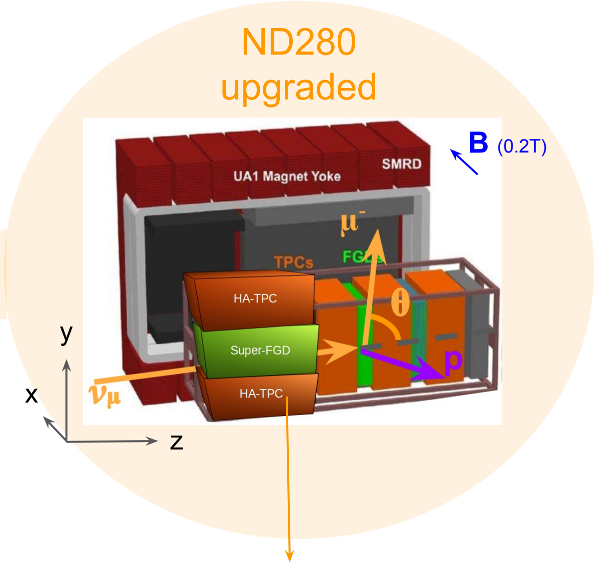


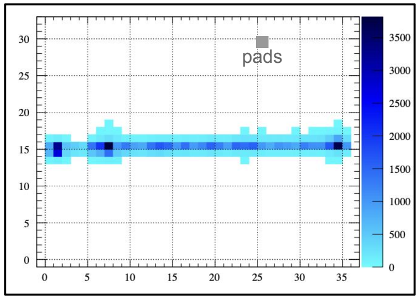
The T2K experiment and its Near Detector

ND280 and its HA-TPCs

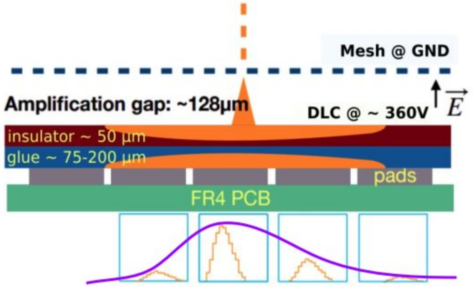
- 50 kt water Cherenkov detector
- 13 000 PMTs

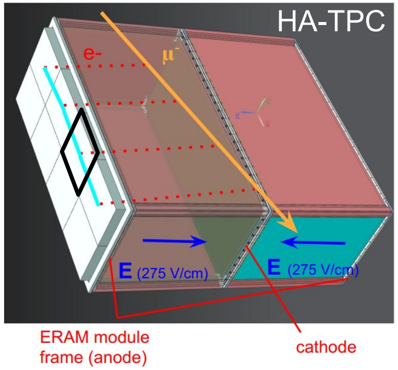






Encapsulated Resistive Anode MicroMegas (ERAM)





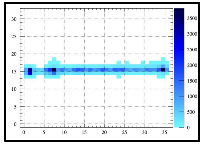
Overview

- T2K and its Near Detector ND280
- **CNN for track reconstruction (the theory)**
- **CNN for track reconstruction (in practice)**
- Results on momentum reconstruction
- Results on Particle IDentification

CNN for track reconstruction (the theory)

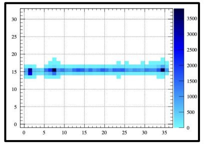
Convolutional Neural Network

- Initial idea: use a CNN to extract particle momentum and PID from the detector "images" (assuming track ID and isolation)



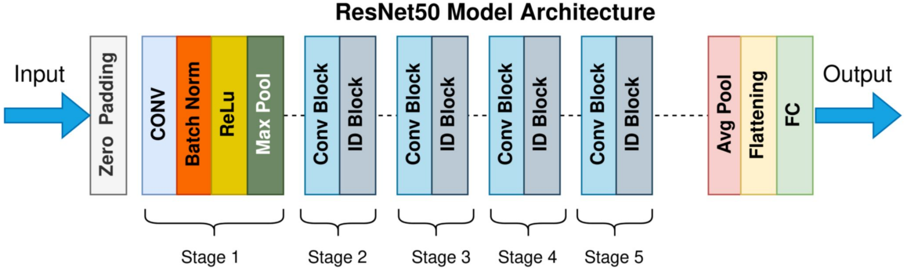
CNN for track reconstruction (the theory)

Convolutional Neural Network



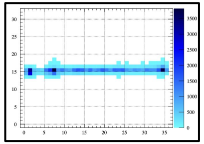
- Initial idea: use a CNN to extract particle momentum and PID from the detector "images" (assuming track ID and isolation)

- CNN choice: ResNet50



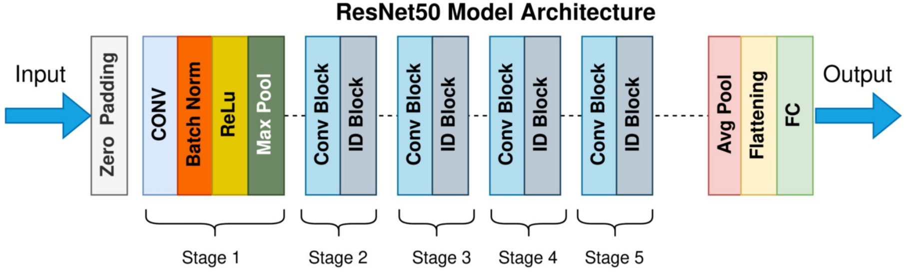
CNN for track reconstruction (the theory)

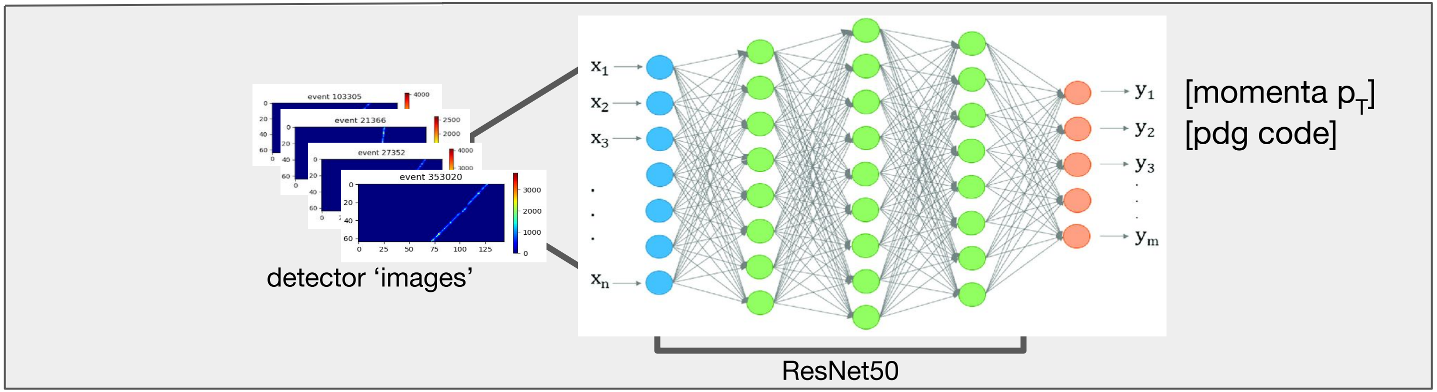
Convolutional Neural Network



- Initial idea: use a CNN to extract particle momentum and PID from the detector "images" (assuming track ID and isolation)

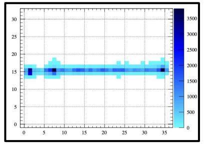
- CNN choice: ResNet50





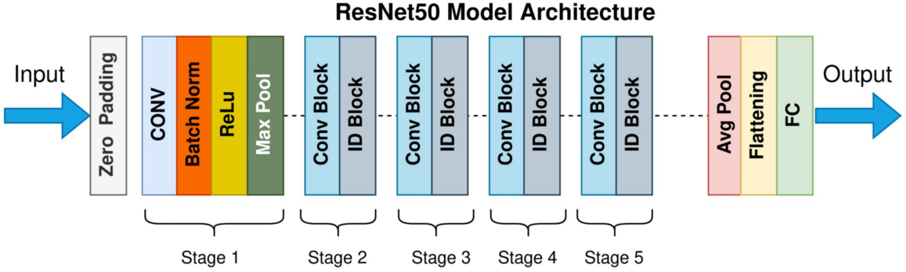
CNN for track reconstruction (the theory)

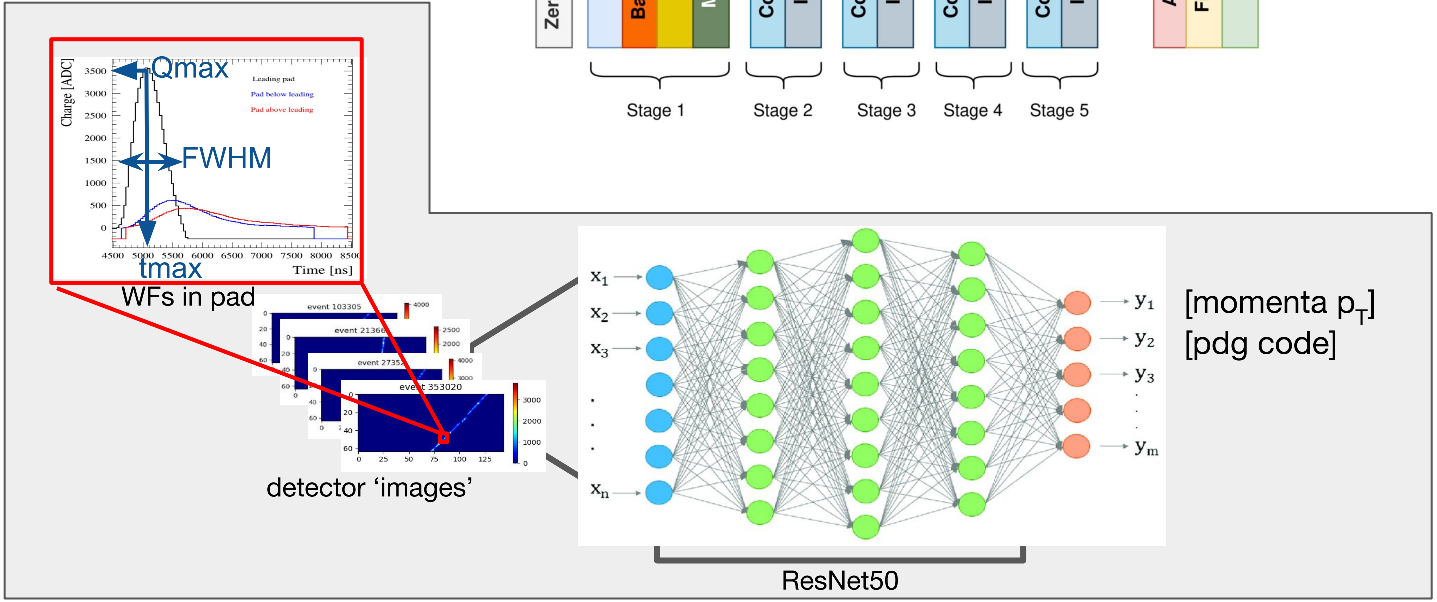
Convolutional Neural Network



- Initial idea: use a CNN to extract particle momentum and PID from the detector “images” (assuming track ID and isolation)

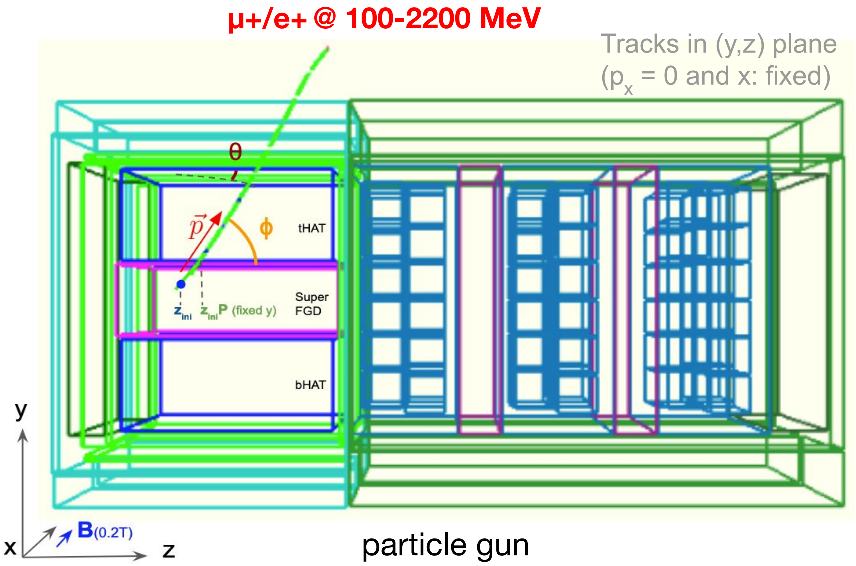
- CNN choice: ResNet50





CNN for track reconstruction (in practice)

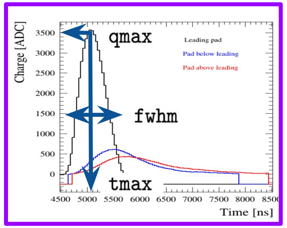
Simulation used

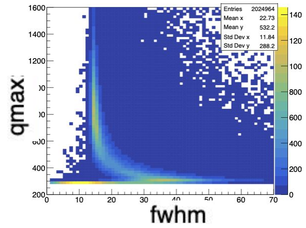


200 000 to 800 000 events generated

CNN for track reconstruction (in practice)

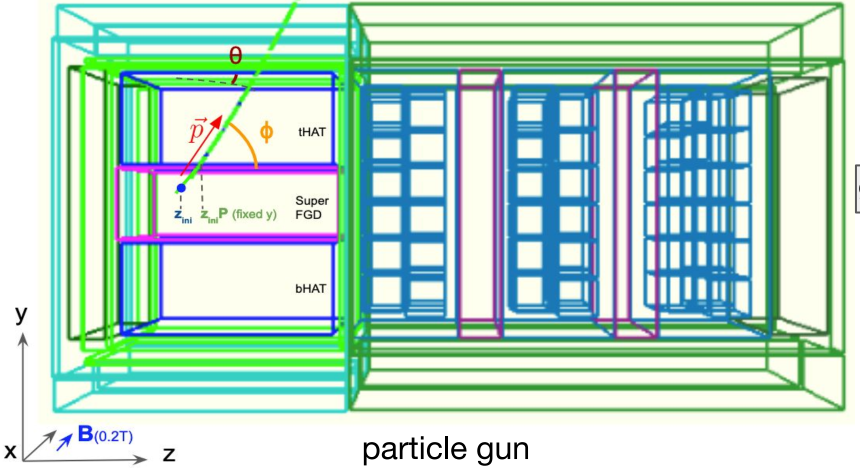
Simulation used





μ+/e+ @ 100-2200 MeV

Tracks in (y,z) plane
(p_x = 0 and x: fixed)



particle gun

gives

inputs
what is fed to the NN

Saved param	Description
qmax	WF maximum in the pad
tmax	Time associated to qmax in the pad
fwhm	FWHM of the WF in the pad
row	row number of the pad in the fem
col	col number of the pad in the fem
fem	FEM (Front-End Mezzanine) i.e. ERAM number
true_event	Event number
true_nhits	Number of hits of the track
true_pdg	PDG code of the particle (-13 for μ^+, -11 for e^+)
true_hat_start	Position (3D) at the HAT entrance
true_mom	Track initial 4-momentum

"measured info"
information in each pads from the detector response

true information
used to train the network

targets
what we want to predict with the network

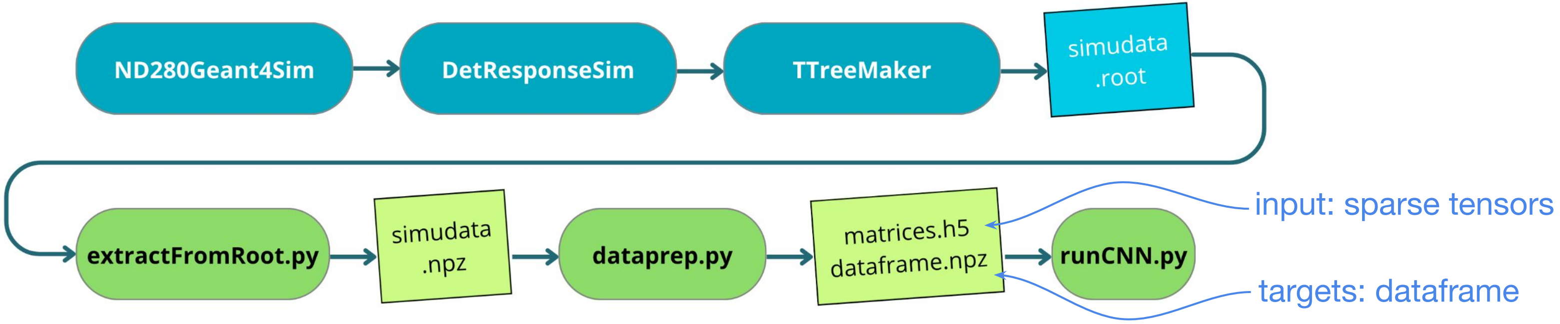
200 000 to 800 000 events generated

CNN for track reconstruction (in practice)

python™
PyTorch
Pandas

Pipeline developped

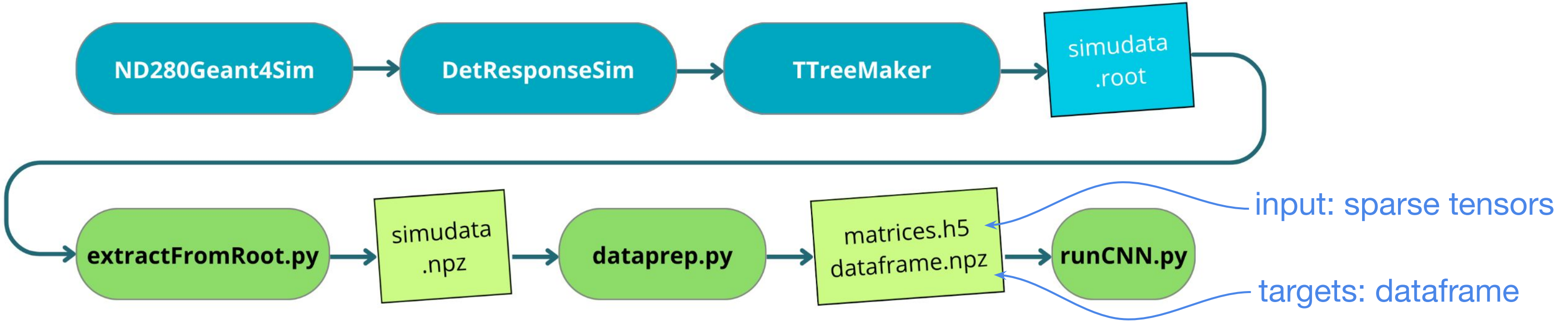
(1) convert ROOT data into readable python data



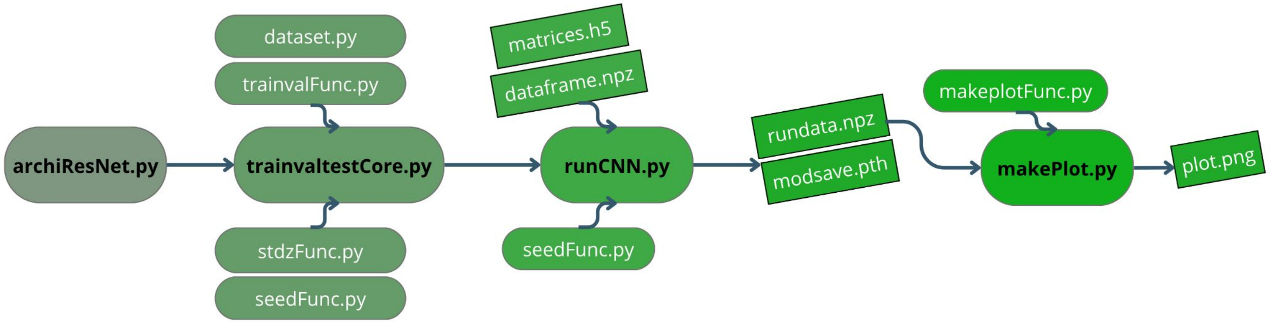
CNN for track reconstruction (in practice)

Pipeline developped

(1) convert ROOT data into readable python data



(2) run the CNN



CNN for track reconstruction (in practice)

Hyperparameters

- Loss function: **Mean Square Error:**

```
torch.nn.MSELoss()
```

$$\underset{\text{or 'loss'}}{cost(w,b)} = \frac{1}{N}\sum_{j=1}^{N}\left[y^j_{pred}(w,b) - y^j_{true}\right]^2$$

- Optimizer: **Adam = variant of Stochastic Gradient Descent:**

```
optimizer = torch.optim.Adam
```

- Hyper-parameters:

-batch size: 64
-epochs: usually in [20,50] depending on data size
-initial learning rate = 0.001 or 0.01
-learning rate patience = 3

(no HPO performed, just hand-tuned)

- Regularisation: target standardisation and dropout (0.5)

- Train/validation/test split: 70%/15%/15%

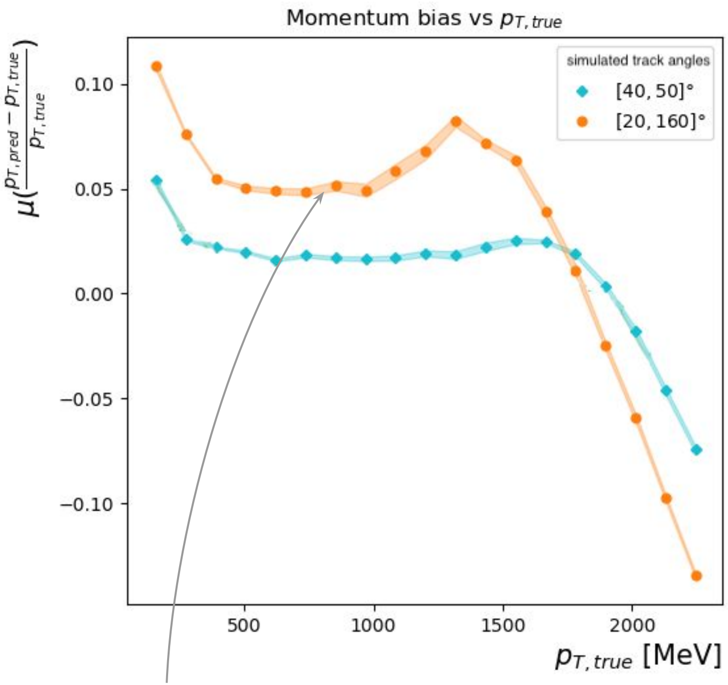
Overview

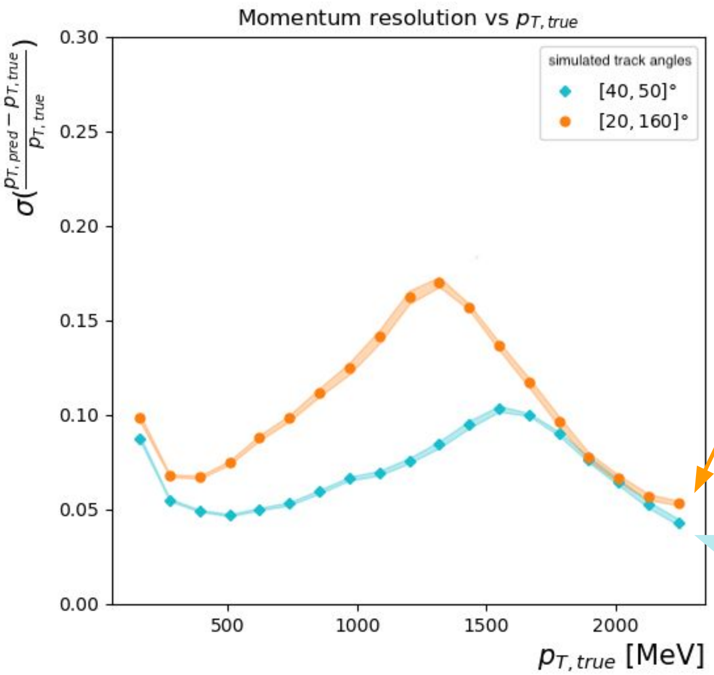
- T2K and its Near Detector ND280
- CNN for track reconstruction (the theory)
- CNN for track reconstruction (in practice)
- **Results on momentum reconstruction**
- **Results on Particle IDentification**

Results on momentum reconstruction

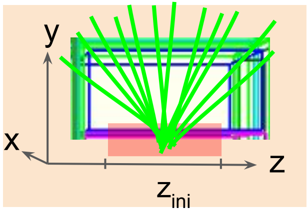
Training on different angular range

larger angle range leads to worst resolution
(more complexe/diverse data)

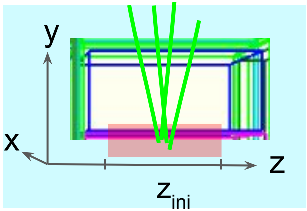




[20,160]° ; 560k events



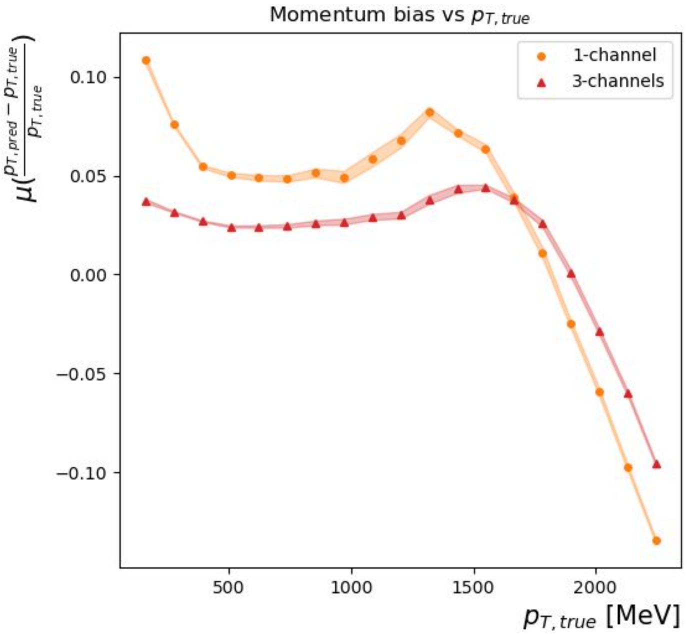
[40,50]° ; 350k events

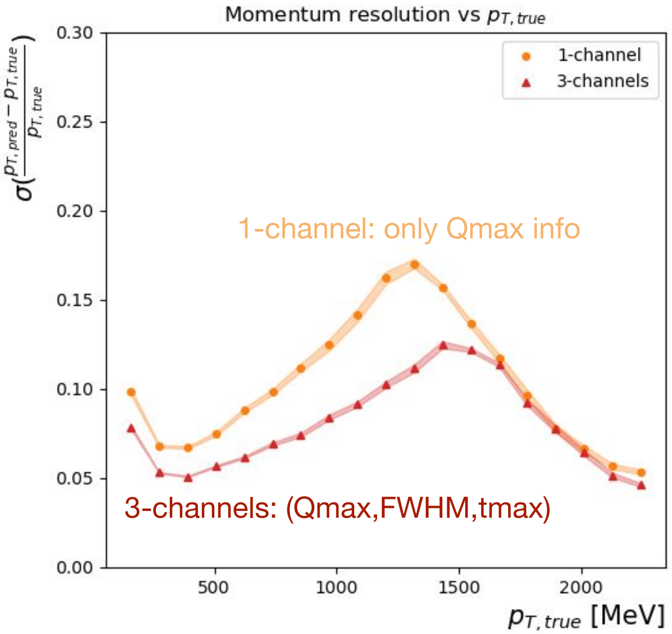


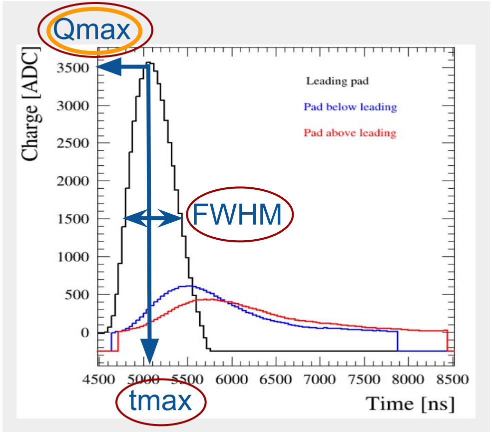
shaded: uncertainties from gaussian fits which extract μ,σ

Results on momentum reconstruction

Using 3 channels



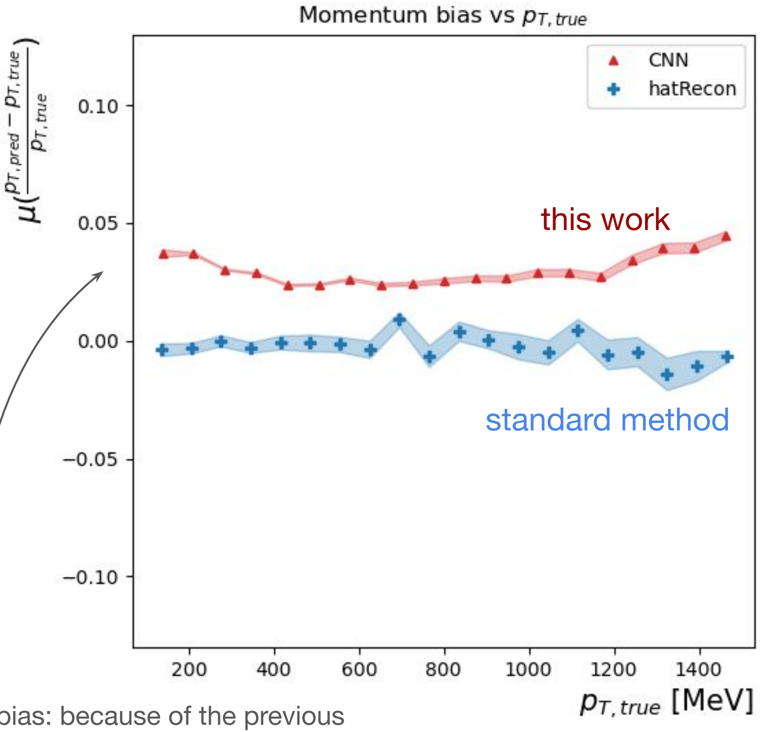


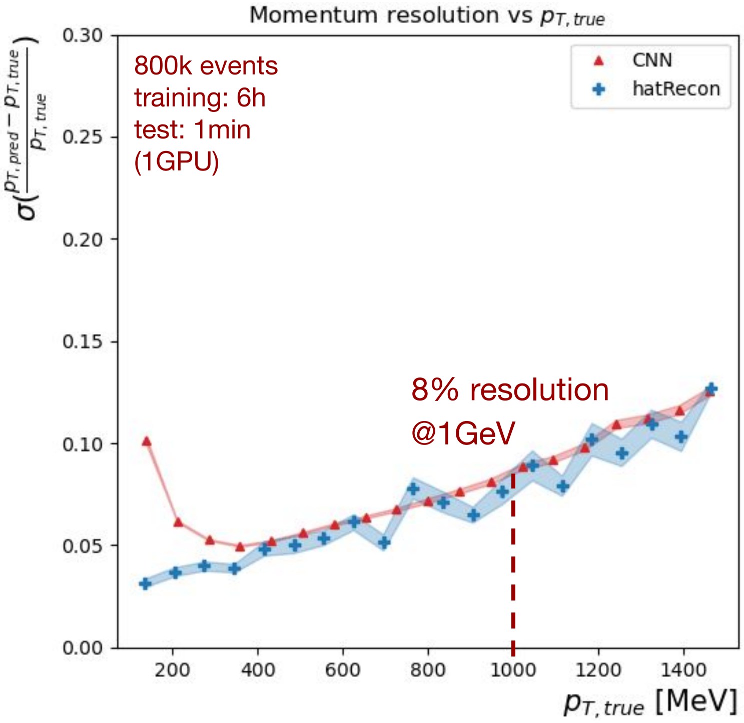


Results on momentum reconstruction

Compared to the standard reconstruction method

trained on [100-2200] MeV range **but** tested on [100-1500] MeV to get rid of border effect



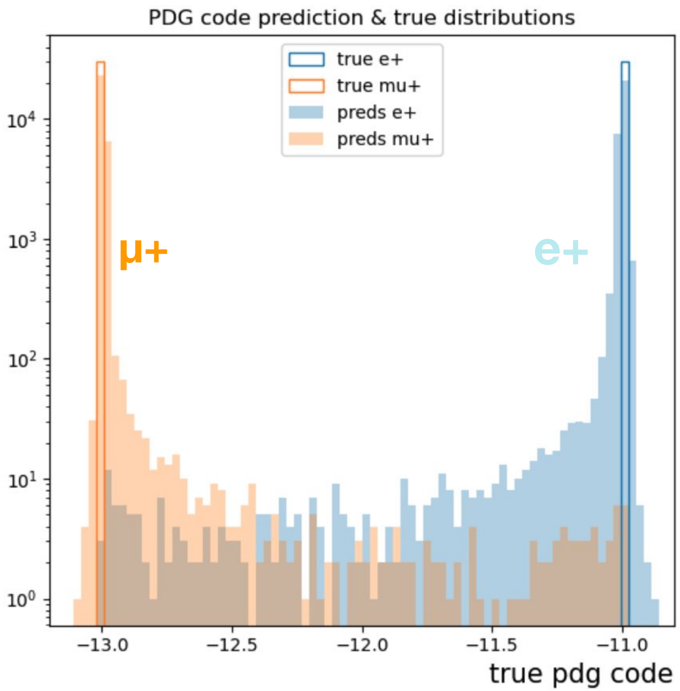


positive bias: because of the previous shape + momenta cut for testing

Results on Particle IDentification

(with a similar regression task, and not a classification one: prediction of the PDG code of the particles i.e. -11 or -13)

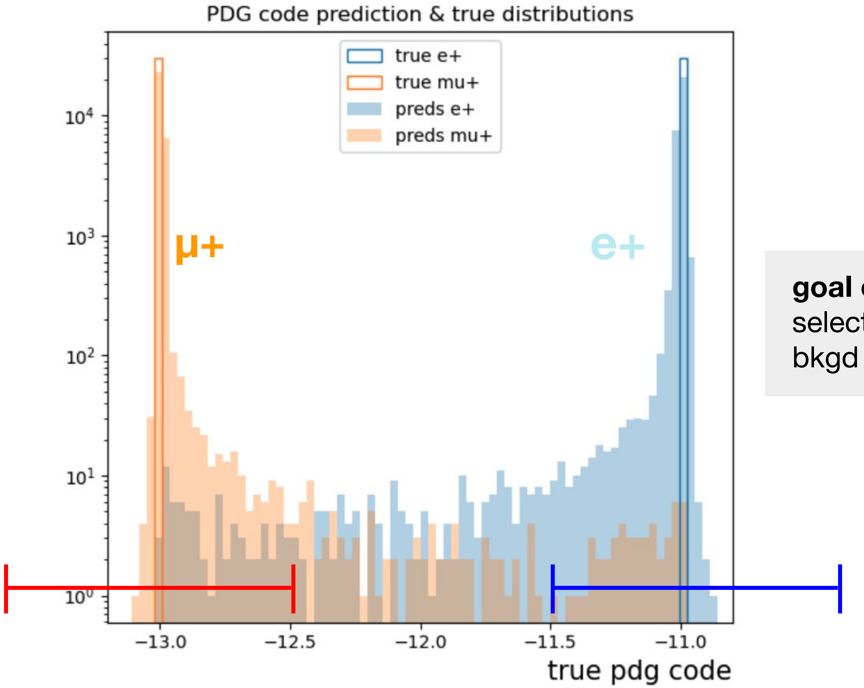
Prediction distributions



Results on Particle IDentification

(with a similar regression task, and not a classification one: prediction of the PDG code of the particles i.e. -11 or -13)

Prediction distributions



goal of PID:
select the ν_e bkgd

$$\text{for e+ selection:} \quad -11.5 < \text{pdg\_code} < -10.5$$

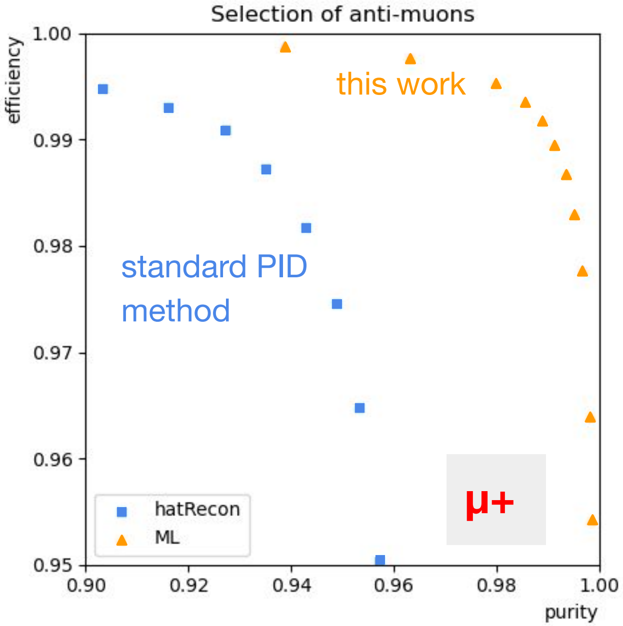
$$\text{for μ+ selection:} \quad -13.5 < \text{pdg\_code} < -11.5$$

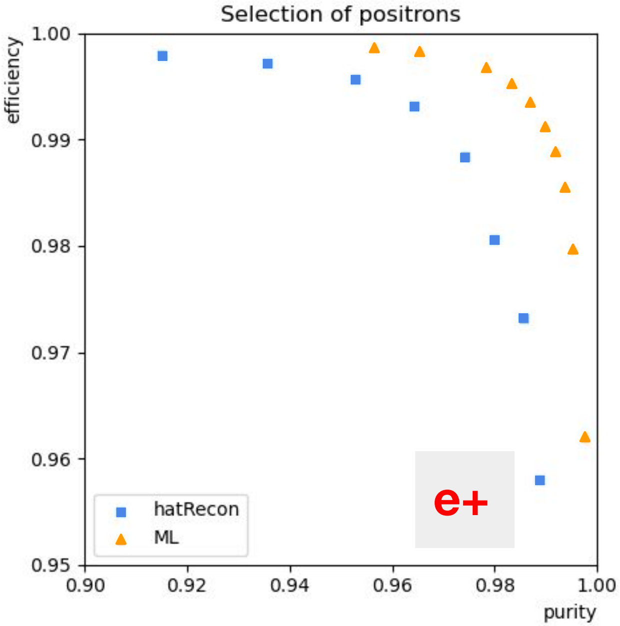
Results on Particle IDentification

Selection performance

$$\text{eff} = N_i^{\text{selected}} / N_i^{\text{generated}}$$

$$\text{pur} = N_i^{\text{selected}} / \left(N_i^{\text{selected}} + N_j^{\text{selected}} \right)$$





Summary

Simulated data from T2K ND280 HA-TPC have been reconstructed with a CNN

- we have demonstrated the **feasibility of momentum reconstruction**

  ↳ we found a momentum resolution very similar to the standard reconstruction algorithm

  ↳ **8% momentun resolution at 1 GeV**

- we have shown **better PID performance** than the truncated mean method in use

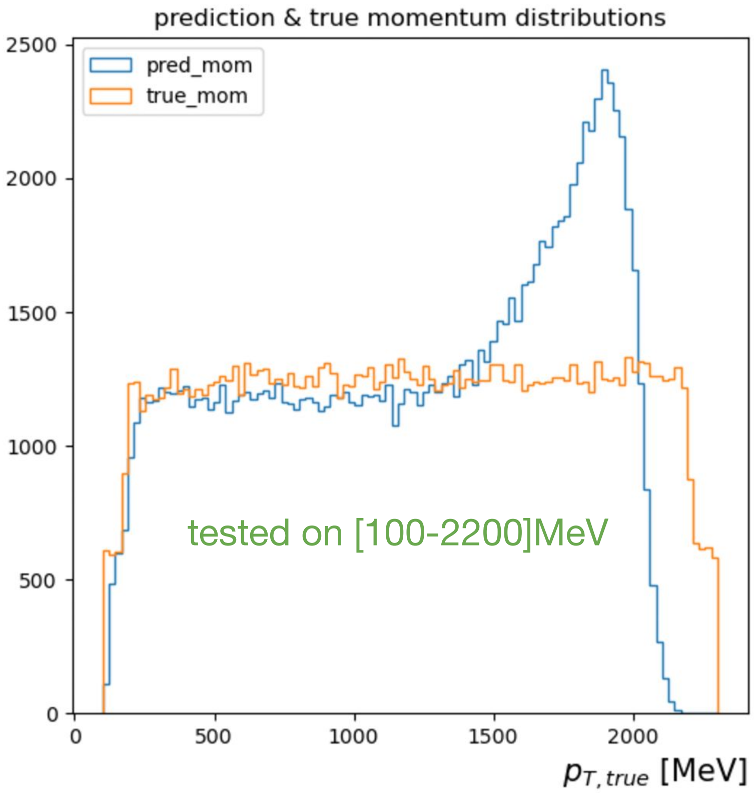
- still many ways to improve: study edge effect, more features, different NN…

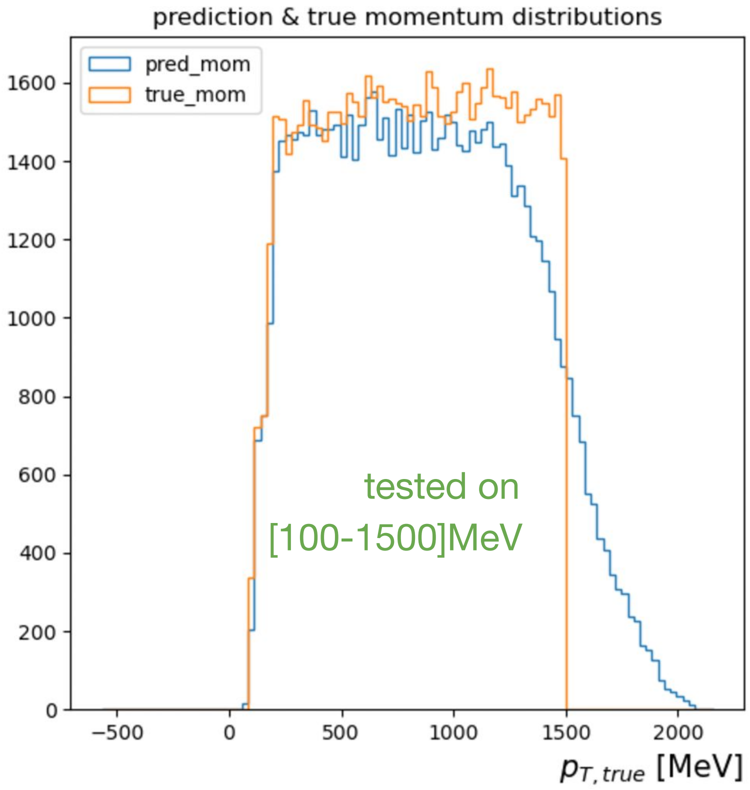
Back-up

Back-up

Edge effect

both trained on [100-2200]MeV

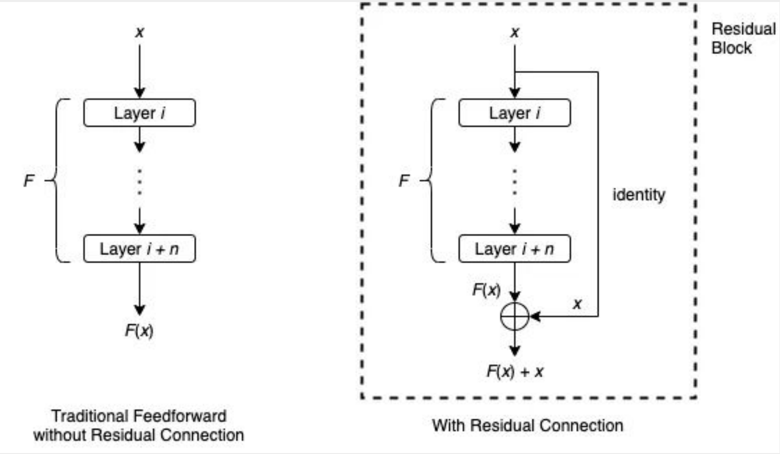




Back-up

ResNet50

Residual NN



Back-up

ResNet50

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9