Machine learning applications in particle physics

Universität Zürich[UZH]

D. Lancierini (davide.lancierini@uzh.ch)

S. Liechti (sascha.liechti2@uzh.ch), P. Owen (patrick.owen@cern.ch)

Group Nicola Serra

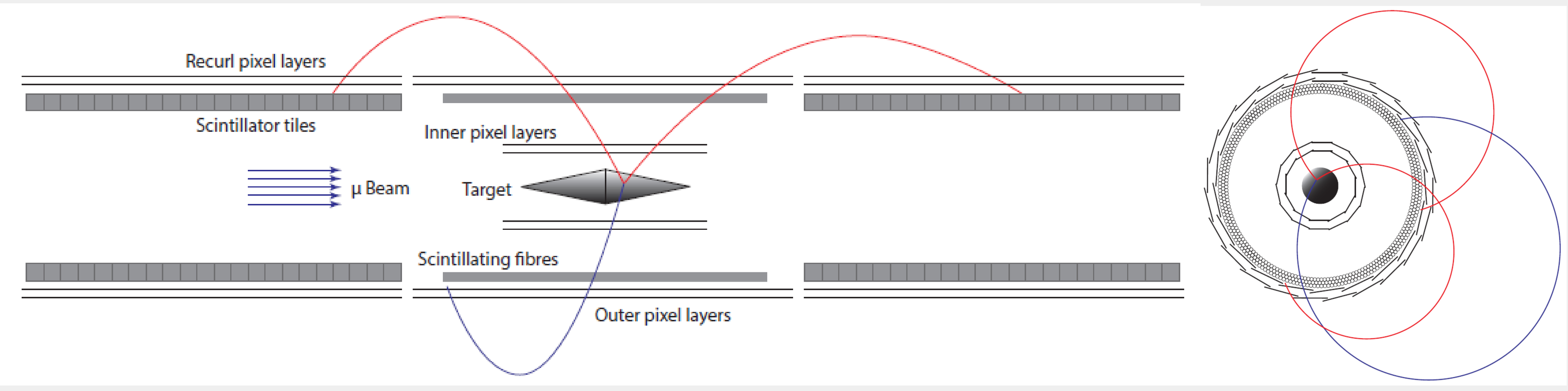
Neural Networks

Neural networks are extremely versatile functions approximators. They are trained by minimising a loss function $\mathcal{L}$ with respect to its parameters.

They are known as SUPERVISED NETWORKS if the network has access to the ground truth while training, and the $\mathcal{L}$ is a function of the labels. (classification, regression)

If they minimise a loss $\mathcal{L}$ that doesn't contain the labels of the training set they are called UNSUPERVISED NETWORKS (autoencoders, generative models)

Search for charged LFV at Mu3e



- Search for charged lepton flavor violation (i.e. $\mu^- \to e^- e^+ e^-$)
- Electrons produced are deflected back into the detector and then stopped in the outer stations of the detector (Energy measurement)
- Current track reconstruction uses a χ^2 method
- If resulting particles have a low pseudorapidity, they get deflected back into the central station

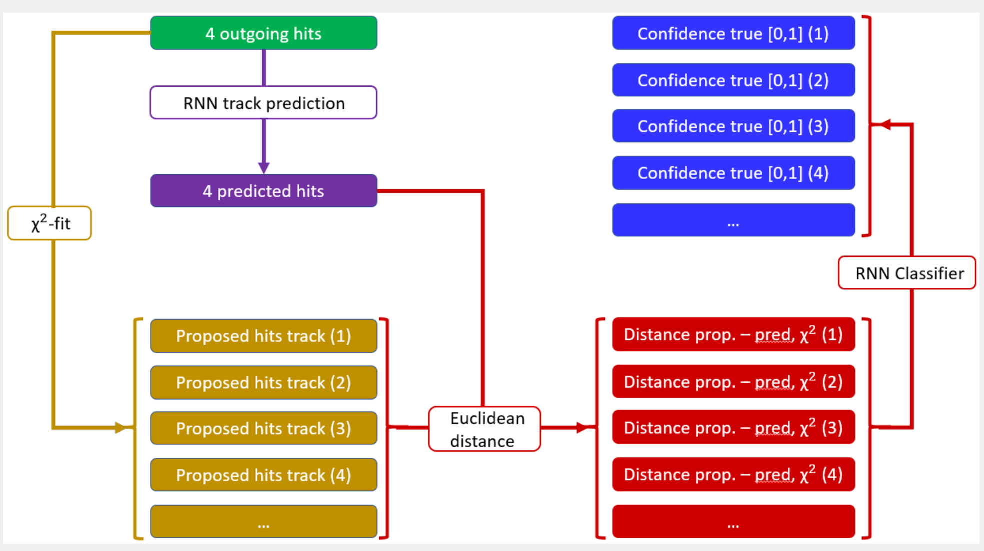
▷ There they pass the central station many times over, creating additional unwanted hits

▷ χ^2 method reconstructs the right path only in around **52%** of the cases

Recurrent Neural Networks (RNN)

1. RNN:

- Produced particle leaving central station creates 4 hits
- Onedirectional RNN predicts next 4 hits when the particle reenters the central station.



2. RNN:

- Introduce a cut for the χ^2 value

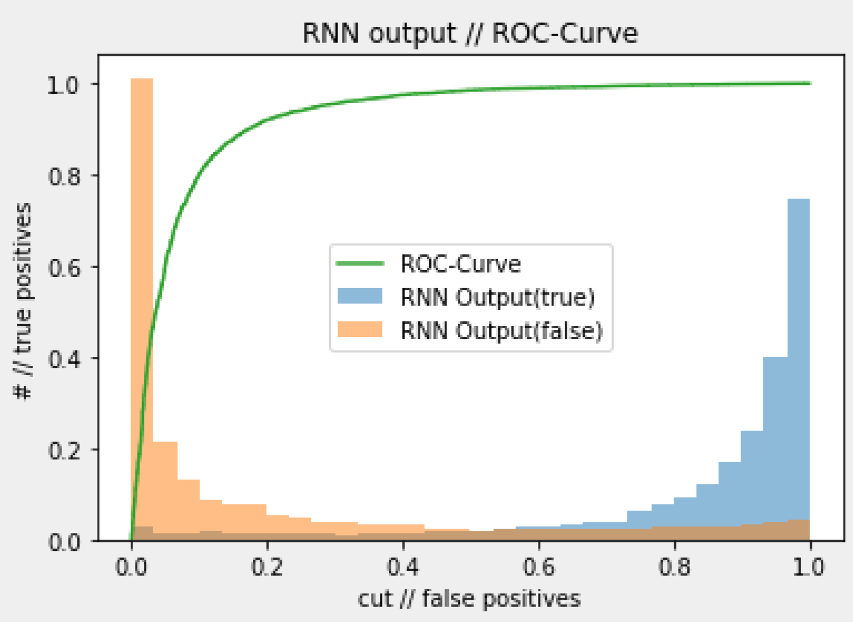
▷ Selection of possible tracks

- Euclidean distance between the predicted and the preselected track(coordinate-wise)
- Euclidean distance and χ^2 value is given to a second bidirectional RNN

▷ Returns confidence **[0, 1]** for each track

Results

Model	Accuracy (cut at **0.5**)	ROC-AUC
Best χ^2	**52.01%**	/
RNN	**87.63%**	**0.93**



▷ RNN's are viable solution to this problem

▷ Huge gain in accuracy compared to traditional methods

Future Prospects

Possible improvements:

- Use a fully connected neural network

▷ Both RNN's are connected and train as a unit

- Allow arbitrary track length (dynamic RNN)

▷ Also useful for this experiment as particles often recurl several times

End Goal:

Replace the tracking completely through RNNs

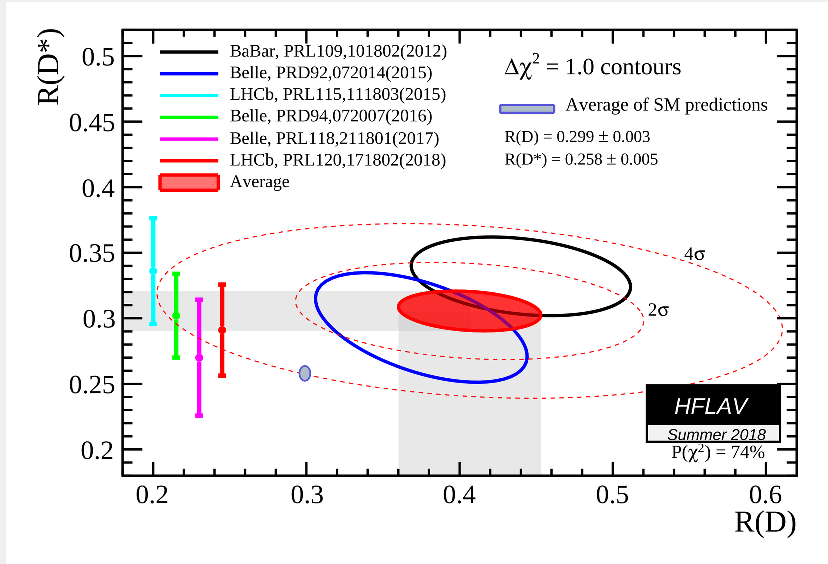
▷ Huge gains in accuracy

▷ Big gains in speed over the χ^2.

B anomalies in semileptonic B-decays

Experimental results show **coherent**, albeit singularly inconclusive, hints of lepton flavour universality violation (LFUV) in semi-leptonic B decays

$$R(D^{(*)}) = \frac{\mathbf{B} \to \mathbf{D}^{(*)}\tau\nu}{\mathbf{B} \to \mathbf{D}^{(*)}\mu\nu}$$



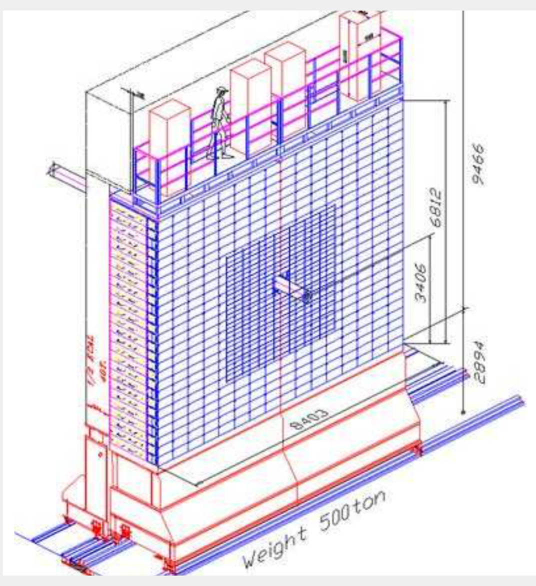
The $R(D^{(*)})$ observable is **theoretically** clean, but **experimentally** difficult due to the missing neutrino(s) in the final state.

Large sample of Monte Carlo data is needed in order to perform the $R(D)$ analysis at LHCb.

Simulating LHCb Calorimeter response with GANs

Monte Carlo simulation of LHCb HCAL for trigger efficiency estimation purposes is time consuming.

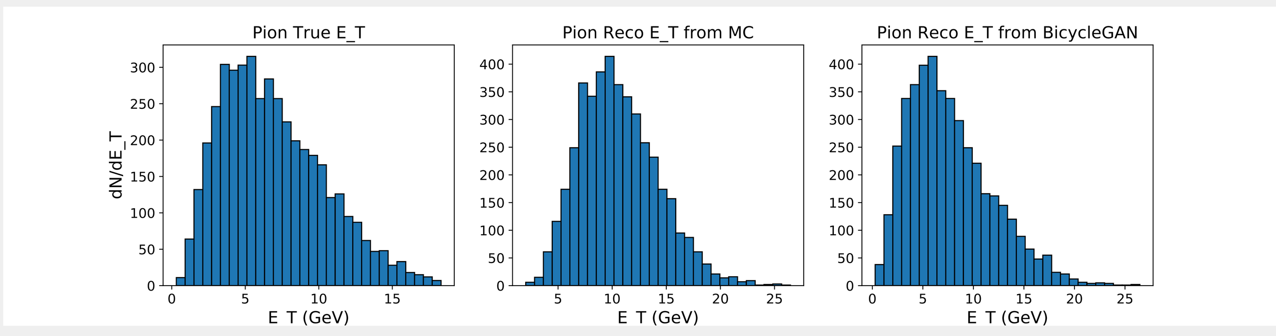
- ▶ HCAL response can be thought as an image: each pixel corresponds to a cell and it's value to the transverse energy recorded
- ▶ Generative Adversarial Networks (DCGAN) can be trained to reproduce distributions of multidimensional datasets



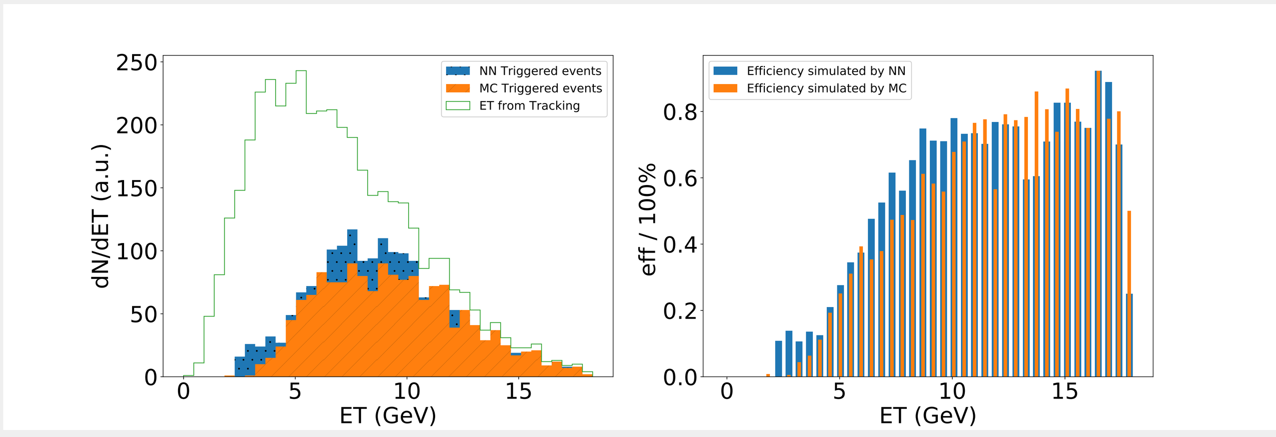
Bicycle GANs as a tool for fast simulation

BicycleGANs map onto each other different distributions of images. They are used to connect sets of images containing the Monte Carlo truth with the corresponding images of the HCAL response.

- ▶ Summing over the energy recorded by each cell:



- ▶ Our aim is to simulate L0 Hadron trigger (fast! $\mathcal{O} \sim 10^{-3}s$ per event):



Future prospects

- ▶ Extend the network capabilities to simulation of HCAL response for a full event and verify its performance, fine tune the obtained network on data.
- ▶ Integrate in LHCb simulation software and extend to future analyses in order to cope with the increased luminosity in Upgrade I
- ▶ Use the gained expertise to extend the use of generative models to different datasets

http://www.physik.uzh.ch/en/researcharea/lhcb.html