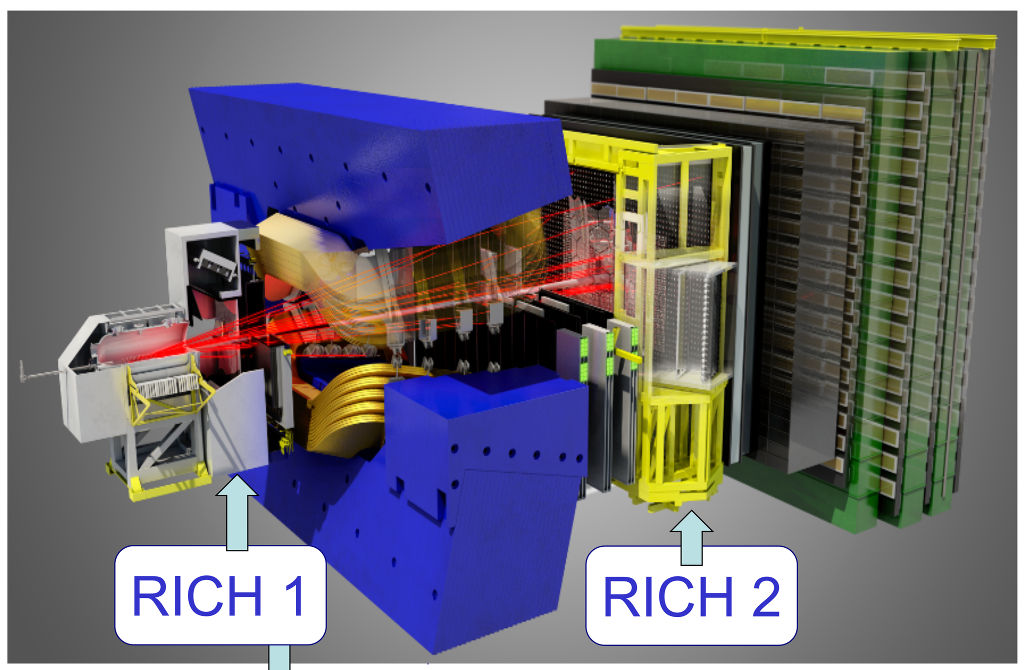
Deep Learning Particle Identification in LHCb RICH

Michele Piero Blago (CERN & University of Cambridge)

Particle Identification in LHCb

The LHCb (Large Hadron Collider beauty) experiment is one of the four major experiments at the LHC (Large Hadron Collider) at CERN (European Organization for Nuclear Research). A schematic is shown on the right. LHCb is dedicated to high-sensitivity searches of Charge-Parity violations in beauty hadron decays and further high-precision heavy-flavour studies [1].

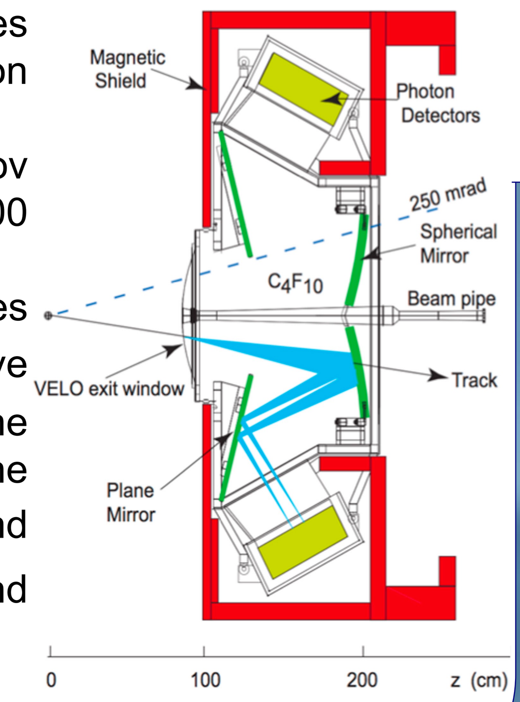
The two Ring Imaging Cherenkov (RICH) detectors in LHCb are responsible for particle identification. Charged particles originating from the LHC collisions traverse the detector enclosures which are filled with gas. Cherenkov radiation is emitted under a characteristic angle θ_{Ckv} relative to the track as a function of the particle's velocity.



LHCb is currently undergoing a major upgrade. After its completion, the RICH detectors will use multi-anode photomultiplier tubes (MaPMTs) for photon detection, collecting data at the LHC collision rate of 40 MHz.

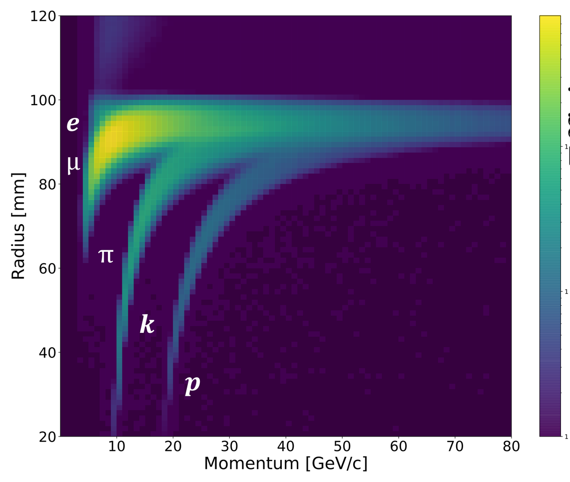
C_4F_{10} (RICH 1) and CF_4 (RICH 2) gas is used as Cherenkov radiators. Particles are identified in an energy range between 2 – 100 GeV/c.

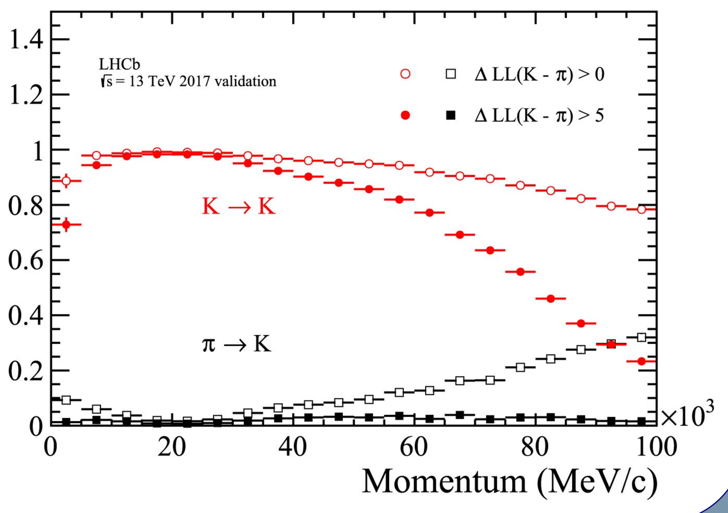
Particles generate Cherenkov photons in the radiators for velocities above the Cherenkov threshold $\beta > \frac{1}{n}$, depending on the refractive index of the radiator. For high momenta, all particles converge at the same Cherenkov radius. The particle velocity is determined via the Cherenkov angle: $\cos\theta_{Ckv} = \frac{1}{n(\lambda)\beta}$. Using the momentum and trajectory curvature information from the tracking, the mass and hence particle species can be determined.



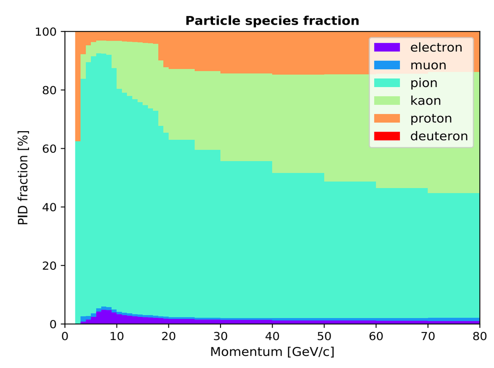
A standard metric for particle identification performance is comparing the kaon identification efficiency (red) and the pion misidentification rate, which represents the most abundant particle species [2].

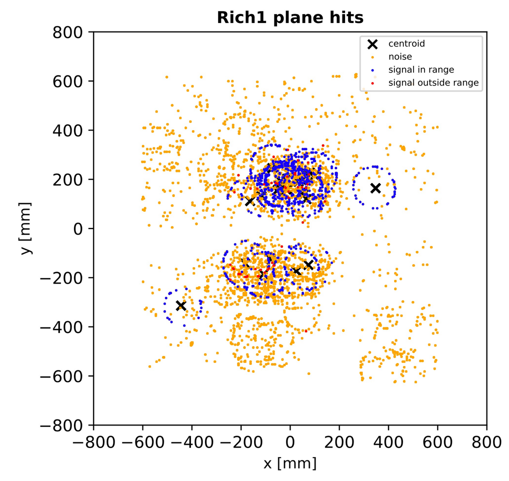
The current algorithm performs very well, however, its complexity is not well applicable to parallelisation on multi-core computing architectures, as foreseen for the LHCb computing infrastructure [3,4].





Simulated RICH data





The primary goal of the RICH detectors is to distinguish between charged hadrons, in particular pions, kaons, and protons. However, electrons, muons, and deuterons also generate a signal in the detector (top).

For the neural network training, LHC Run 3 B_s^0 and $\overline{B_s^0}$ signal events were generated using LHCb simulation and reconstruction software. This allows assessing the performance for the upcoming run period and training the network with labelled data.

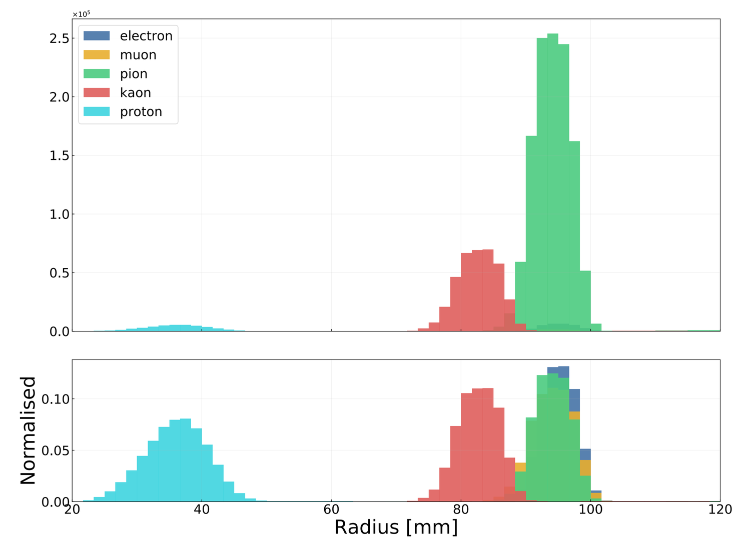
A typical event with five primary vertices is shown (bottom). The track centres of reconstructable tracks, extrapolated to the photon detection plane, are marked by a cross, their signal in blue and the background in orange.

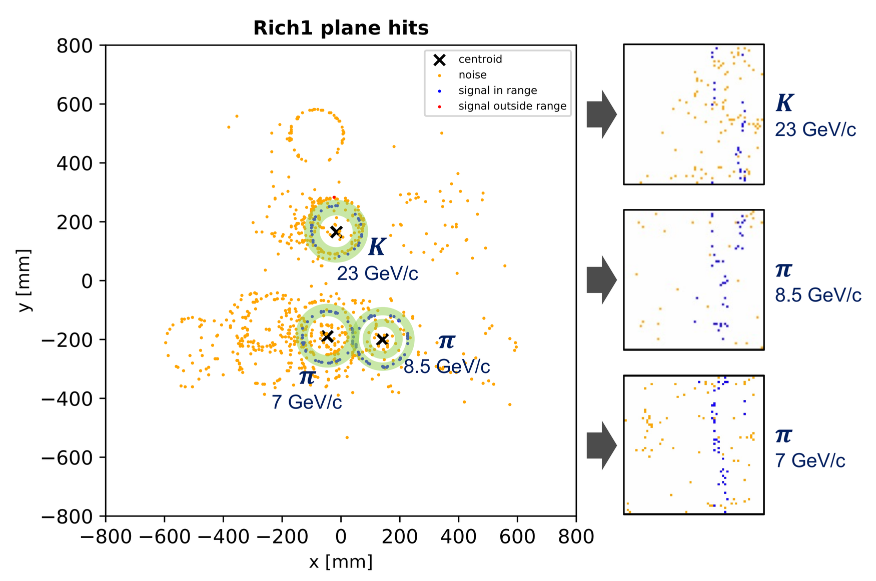
Data preparation and feature engineering

The LHC collision event data is processed to suit as input to the network. It is separated into momentum ranges, with 1 GeV/c bins up to 20 GeV/c and10 GeV/c bins above. For each momentum range the network is trained and assessed separately.

Around each extrapolated track centre, a radius range is defined. The hits in this range are polar-transformed to 64 x 64 images, which represent the input to the neural network.

In the figures below, the radius ranges for the momentum range 19 - 20 GeV/c are shown (left) and the polar transformation demonstrated for a RICH 1 event with a single primary vertex (right). The colour separation of signal and background is for the benefit of the reader, the network receives monochrome images.

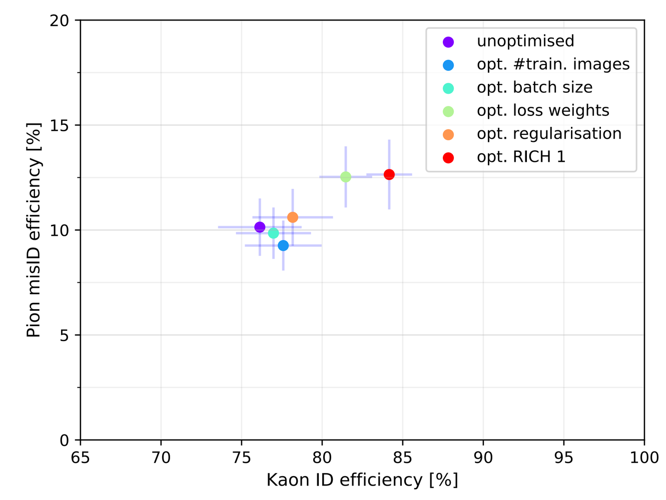


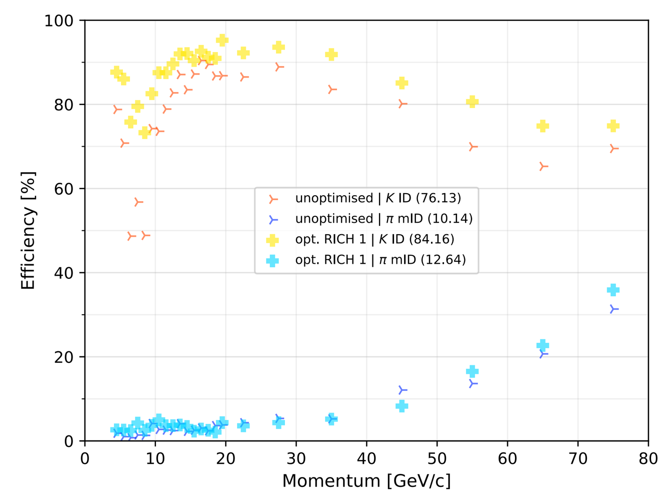


Hyperparameter optimisation

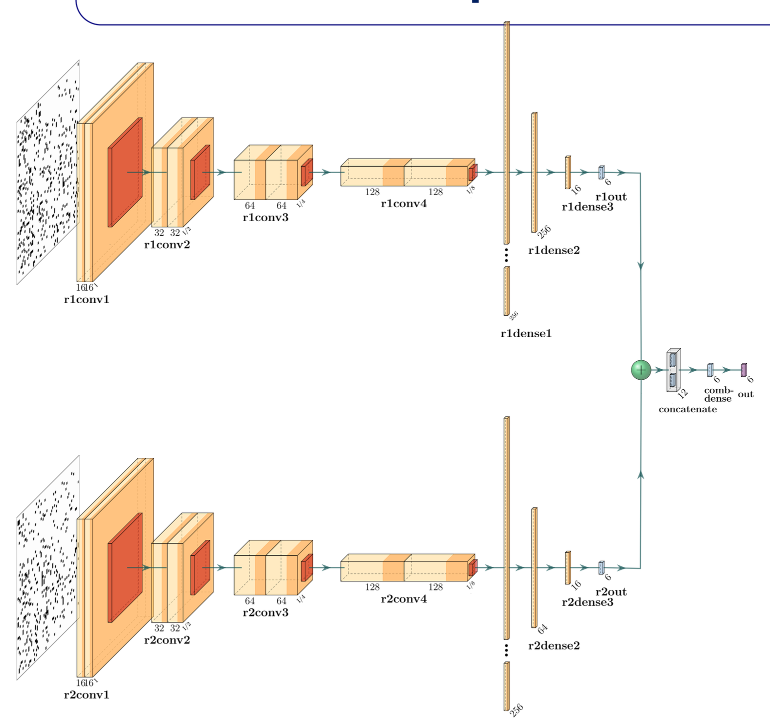
A simple neural network is used as a starting point, using exclusively RICH 1 input images [5]. The network consists of convolutional layers with image dimensions decreasing by max pooling and increasing number of filters (32, 64, 128), followed by two fully connected layers of 256 and 16 nodes, respectively, using 25 % dropout and early stopping for regularisation. *Leaky ReLU* is used as activation function, apart from the output layer where *softmax* is used for the six nodes of the output layer, one for each particle type: **electron**, **muon**, **kaon**, **proton**, **deuteron**. *Categorical-crossentropy* is used as the loss function, *Adadelta* as optimiser, and *accuracy* as metric.

Several parameters on the network performance are assessed, such as the dataset size, batch size and batch normalisation, loss function weights, additional network layers and inputs, and further regularisation options. The effect of each optimised parameter (top) is compared to combining them in an new version of the network (bottom).



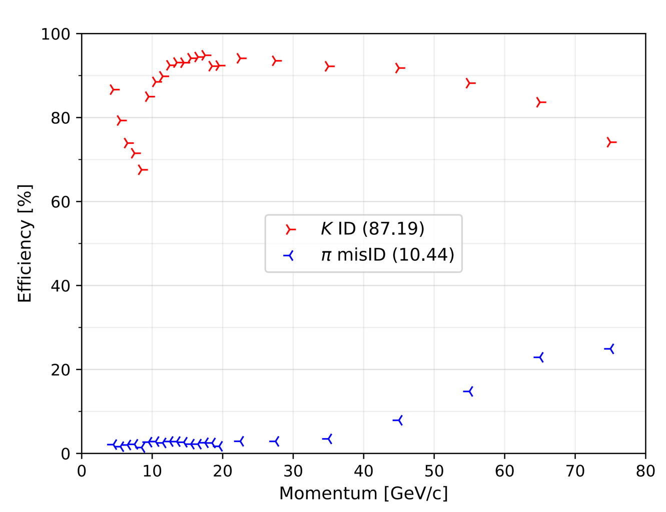


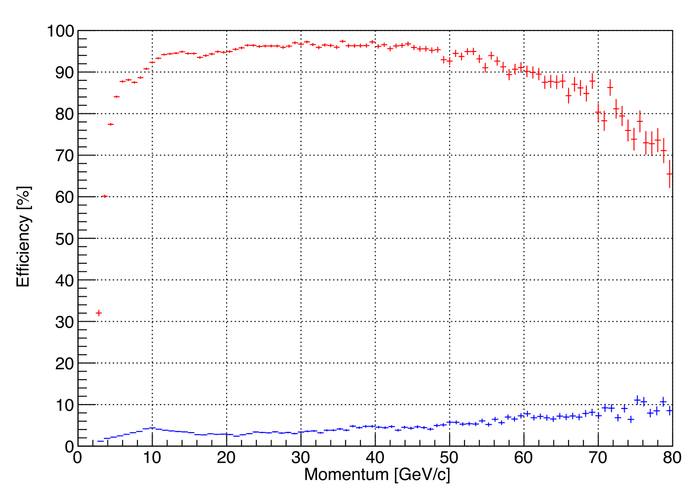
CNN PID performance



The optimised network structure for RICH 1 images is duplicated to include inputs from both RICH detectors. Their output is concatenated and processed by an additional fully connected layer before the final particle ID output (top left).

The performance of the optimised convolutional neural network (bottom left) achieves an average kaon ID efficiency of 87 % with a simultaneous average pion misidentification rate of 10 %. The network performance is comparable to the conventional algorithm (bottom right), with an excellent performance below 50 GeV/c.



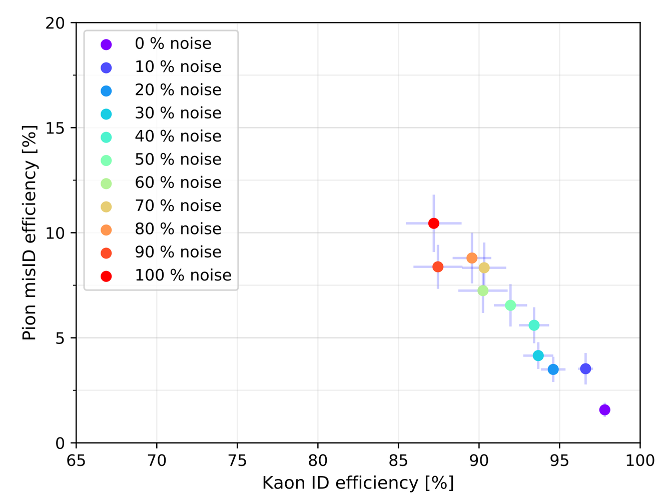


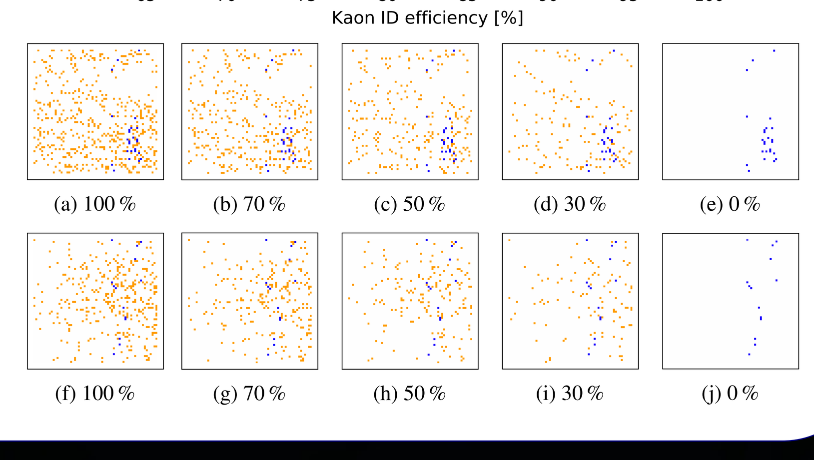
Reducing combinatorial background

In future upgrades of the RICH detectors, the combinatorial background may be reduced by the introduction of timing information [6]. The effect on the network performance has been studied, approximating the effect of reduced combinatorial background by randomly removing a fraction of the hits which do not belong to the signal of a given track.

While the scenario of zero combinatorial background is idealistic, a clear trend towards optimal kaon ID and pion misID is observable, showing the potential of the network to operate in future detector environments.

Details about the presented study can be found in [7].





References

[1] LHCb collaboration. "LHCb PID Upgrade Technical Design Report", CERN-LHCC- 2013-022. November 2013. https://cds.cern.ch/record/1624074

[2] Gambetta, S. et al. "The LHCb RICH detectors: Operations and performance." *Nucl. Instrum. Meth. A*, 2020. doi: 10.1016/j.nima.2019.02.009.

[3] Aaij, R. et al. "Allen: A high-level trigger on GPUs for LHCb." *Comp. and Softw. for big Science*, 4(1):1–11, 2020. doi: 10.1007/s41781-020-00039-7.

[4] LHCb Collaboration. "LHCb Upgrade GPU High Level Trigger Technical Design Report." May 2020. https://cds.cern.ch/record/2717938. CERN-LHCC-2020-006.

[5] Campora Perez, D. and Jones, C. Private communication, 2018.

[6] Keizer, F. "Novel photon timing techniques in the LHCb RICH upgrade programme". https://cds.cern.ch/record/2770576.

[7] Blago, M. "Convolutional neural networks and photonic crystals for particle identification at high energy collider experiments". Doctoral Thesis. October 2021

CERN · LHCb · UNIVERSITY OF CAMBRIDGE

20th International Workshop on Advanced Computing and Analysis Techniques (ACAT) in Physics Research · 29 Nov – 03 Dec 2021