Particle Identification at the CERN NA62 Experiment Using Convolutional Neural Networks

Bob Velghe, for W. Fedorko, M. Yu and the NA62 Collaboration

2021 CAP Virtual Congress
June 7, 2021

NA62 Experiment at the CERN SPS

NA62 primary objective: Measure the $K^+ \to \pi^+ \nu\bar{\nu}$ branching ratio. Excellent probe for new physics in the flavour sector.

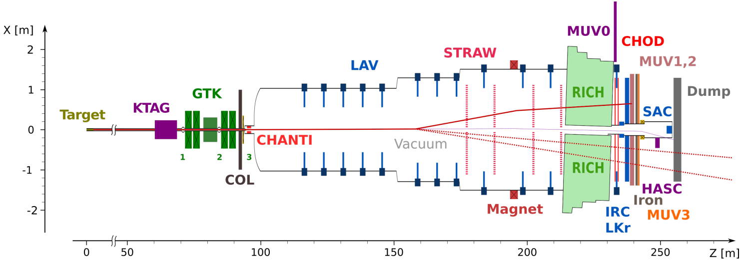
$$\mathcal{B}_{\text{SM.}} = (8.4 \pm 1.0) \times 10^{-11}$$

[A. J. Buras et al, 15']

$$\mathcal{B}_{\text{exp.}} = \left(10.6^{+4.0}_{-3.4}\Big|_{\text{stat.}} \pm 0.9_{\text{syst.}} \right) \times 10^{-11}$$

[NA62 Collaboration, 21']

75 GeV/c mixed hadron beam, Kaon “decay in flight” technique:



For details about the detector, see [NA62 collaboration, 17']

- Identification of K and π
- Multi-track event rejection
- Vetoes for γ and μ, rejection $> 10^7$
- $\mathcal{O}(100 \text{ ps})$ timing for K - π matching
- $m^2_{\text{miss.}} = (P_{K^+} - P_{\pi^+})^2$

$$K^+ \to \mu^+ \nu_\mu \; (\mathcal{B} \approx 0.64)$$
$$\downarrow \text{mis} - \text{id}$$
$$K^+ \to \pi^+ \nu\bar{\nu} \; (\mathcal{B} \approx 10^{-10})$$

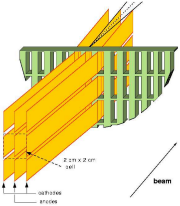
Today's Topic: Calorimetric Particle Identification at NA62

We need an overall μ^+ rejection $> 10^7$ while maintaining adequate π^+ acceptance.

Partially redundant systems: RICH and three **calorimeters**:

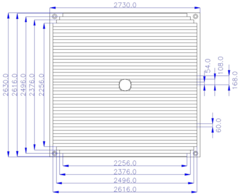
Liquid Krypton Calorimeter (LKr)

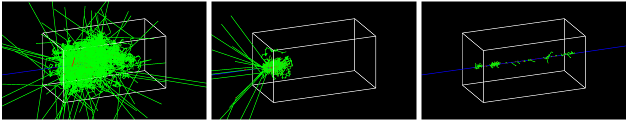
Homogeneous calorimeter, cylindrical shape, $5.3\ \mathrm{m}^2$ front surface, 13248 $2 \times 2 \times 127\ \mathrm{cm}^3$ cells.



MUV1 (and MUV2)

Iron/scintillator sandwich calorimeters, 44 (22) scintillator strips alternately oriented horizontally and vertically. Front surface $2.6 \times 2.6\ \mathrm{m}^2$, about 8 interaction length.





Pions, muons and positrons tracks are selected from common kaon decays modes (**data**):

- $K^+ \to \pi^+\pi^0$ for pions,
- $K^+ \to \mu^+\nu_\mu$ for muons,
- $K^+ \to \pi^0 e^+\nu_e$ for positrons.

Events with a single downstream track matched to a kaon, photon vetos, and

- Pion: π^0 mass reconstruction, RICH, M^2_{miss},
- Muon: RICH, M^2_{miss},
- Positron: π^0 mass reconstruction, RICH, M^2_{miss}.

Training and validation sets:

Data period	Muon	Pion	Positrons
2016A	16 159 022	4 700 336	320 398
2017B	35 610 772	11 111 872	834 078
2018B	31 835 976	8 908 019	670 807

A completely independent test set was kept aside during the development phase (*minimum bias* trigger).

RICH: Ring-imaging Cherenkov

From Energy Deposits to Image Recognition

Current Calo PID algorithm relies on a Boosted Decision Tree (BDT) based on reconstructed quantities such as cluster energies and shapes.

Other approach: Direct correspondence between the readout geometry and a five channels image. No depth information, the shower is projected on a plane.

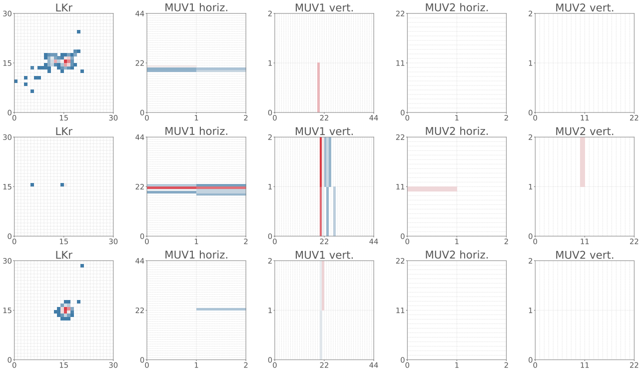


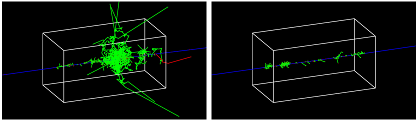
Image preparation:

- **LKr**: 22×22 cells patch centred around the charged track impact point,
- **MUV1**: Full detector, centred around the charged track impact point,
- **MUV2**: Full detector, centred around the charged track impact point.

Processing Pipeline

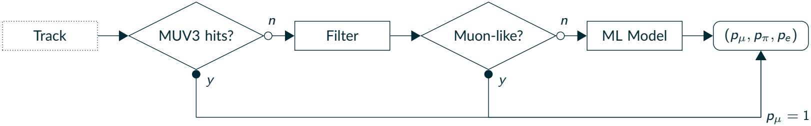
Optimize π/μ separation for single track events in the $15 < P_{\mathrm{trk.}} < 40$ GeV/c range.

- The μ^+ set is dominated by "easy" to distinguish muons. We want the ML algorithm to focus on "hard" cases,



- Similarly, most of the e^+ are easy to identify in the LKr (E/p cut),

$\rightarrow$ Filter those events before training ML models.

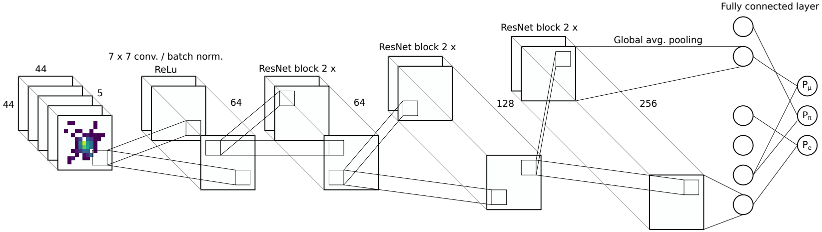


Typical example (2016A):

Event class	Selected	Used for training/validation
μ^+	16 159 022	3 245 995
π^+	4 700 336	3 647 318
e^+	320 398	26 395

Convolutional Neural Network (CNN) Architecture

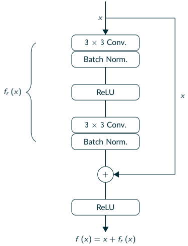
We opted for a simplified ResNet-18 architecture [K. He et al, 15']



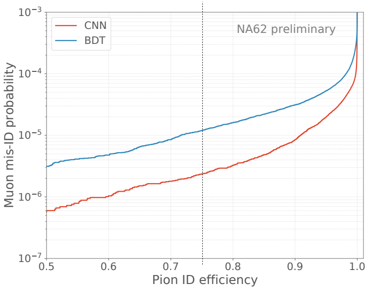
ResNets have been shown to improve the learning performances of deep CNNs. In addition, the shortcut connections help to control the vanishing gradient problem.

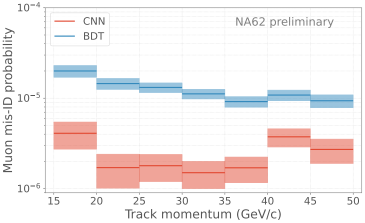
The network was simplified by removing

- the last two residual blocks,
- the 3×3 maximum pooling layer after the first batch normalization layer.



The network performances were evaluated with the unseen test set once the network architecture frozen.

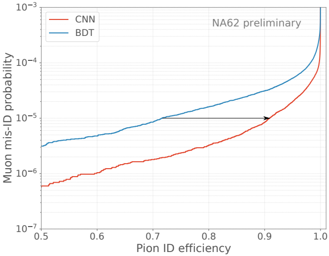




The CNN was trained on tracks in $15 < P_{\mathrm{track}} < 40\ \mathrm{GeV}/c$ range but validated over the $15 < P_{\mathrm{track}} < 50\ \mathrm{GeV}/c$ range.

Training a dedicated CNN per momentum bin was attempted but didn't improve the performances.

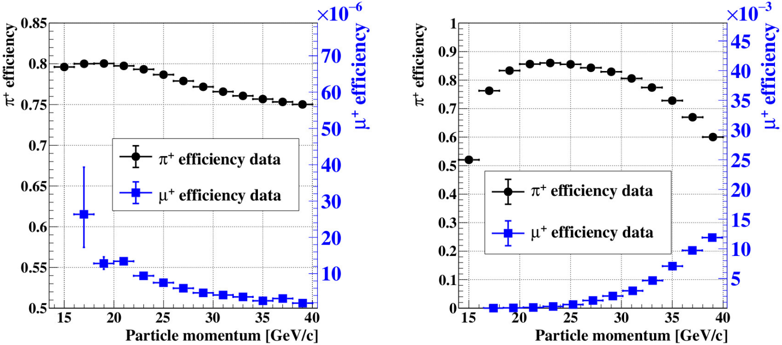
- The NA62 PID system was designed to have a μ^+ mis-id probability $< 10^{-7}$ for an overall pion acceptance around 80%.
  Complementary sub-systems: RICH $\rightarrow \mathcal{O}\left(10^{-3}\right)$ and Calorimeters $\rightarrow \mathcal{O}\left(10^{-5}\right)$.
- This work focuses on the **calorimeters**: A Convolutional Neural Network trained on **raw hits** data outperforms the current BDT.
- π^+ ID efficiency increases from 72% to 92% while keeping the μ^+ mis-id probability at the 10^{-5} level.



- The algorithm is currently being implemented into the NA62 software framework (C++, `libtorch`).


FORTAL

Particle Identification Performances



Performance of the π^+ identification using calorimeters (left) and RICH (right) measured on 2017 data.