CNN-based event filtering in LAr detectors using PMTs

Marta Babicz\textsuperscript{1,2}, Saul Alonso Monsalve\textsuperscript{1}, Stephen Dolan\textsuperscript{1}

\textsuperscript{1}CERN

\textsuperscript{2}The H. Niewodniczański Institute of Nuclear Physics

February 04, 2021

Contact: marta.babicz@cern.ch
Within the BNB spill window we expect over three times more cosmic ray backgrounds than neutrino interactions.

We aim to reduce this background using the information we have available from the PMTs.

The output is fed into a Convolutional Neural Network (CNN) to discriminate between cosmics and genuine neutrino interactions.
**Goal:** Reduce the cosmic background in the BNB trigger window using the information we have available from the PMTs.

- Following the ICARUS trigger, PMT signals are considered per pair of PMTs.
- As a position, we take a 3D position of each pair as the point halfway between them.
- We also use the time each pair went above the threshold in the trigger.
- And the first opening of the trigger gate after applying the beam gate coincidence (one time per channel), so first time that channel opened.
- These are then converted into 3D images which are used to train our CNN to separate cosmics from neutrino interactions.
The primary goal of this tool is to act as an offline event filtering tool, reducing the cosmic background prior to further processing using PMT multiplicity and timing information.

We eventually plan to update the tool to also include PMT waveform information and consider each PMT instead of a pair.
**Why CNN?**

- Separating signal events from background events is a well-studied application of machine learning in HEP.
- ICARUS light detection system - densely packed in PMTs - contains enough detail of interactions and lend itself nicely to image recognition techniques.

- CNNs are designed for image recognition tasks.
- Main concept: apply filters to images to extract features.
- We build images using the information we have available from the PMTs.
**NETWORK ARCHITECTURE**

- **Architecture:** 3D ResNet
- **Parameters:** 33,185,473
- **20k neutrinos + 45k** cosmics for training
- **Optimiser:** SGD with learning rate of 0.1 (divided by 10 when error plateaus), momentum of 0.9, and decay of 0.0001
- **Weighted loss function,** with coefficients: class weights: 1.0 (neutrino), 1.4547077197679608 (cosmic)
- **Trained on NVIDIA V100 GPU**
- We plan to investigate other architectures to be able to use multiple opening times per channel
**OVERVIEW**

New input data used in this study:

- The MC data (v09_06_00 icaruscode, no overburden, events in two cryostats):

<table>
<thead>
<tr>
<th>Type</th>
<th>Tot. number of events</th>
<th>Triggered events</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORSIKA</td>
<td>248600</td>
<td>50697</td>
</tr>
<tr>
<td>GENIE BNB (+filter)</td>
<td>50000 (24728 in AV)</td>
<td>24474</td>
</tr>
<tr>
<td>GENIE BNB (no filter)</td>
<td>59700</td>
<td>44775</td>
</tr>
</tbody>
</table>

- In the first two samples only two categories of events were considered neutrinos and cosmic background (*binary classification problem*).
- For the non-filtered neutrino sample and cosmic muon events we distinguish three categories of events (*multi-class classification problem*):
  - two kinds of neutrinos: OAV* and IAV**
  - cosmic background events.

* OAV - neutrinos that are out of Active Volume  ** IAV - neutrinos inside Active Volume
OVERVIEW

- Real cosmic + BNB beam data:

<table>
<thead>
<tr>
<th>Run number:</th>
<th>4642</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of events:</td>
<td>156758</td>
</tr>
<tr>
<td>Triggered events:</td>
<td>95984</td>
</tr>
<tr>
<td>Data taking time:</td>
<td>14 hours not interrupted</td>
</tr>
<tr>
<td>Beam proton intensity:</td>
<td>2.8E12</td>
</tr>
<tr>
<td>Beam rep. rate:</td>
<td>3 Hz</td>
</tr>
<tr>
<td>Active volume:</td>
<td>East (one cryostat)</td>
</tr>
</tbody>
</table>

Expected events:

- For a proton intensity of 1E12 and 5Hz in 24h:
  - 400 beam events + 400 rock muons
  - 6400 cosmics (0.016 cosmics in 1.6 µs spill)

- For run 4642:
  - 360 beam events / 2400 cosmics (156k spills × 0.016 cosmics)
OVERVIEW OF THE PMT PAIR ACTIVITY (ONE CRYOSTAT)

- Better agreement of simulated to real cosmic data - PMTs have been equalised.
- Note: the y-axis is arbitrary in scale (i.e. we don’t think the data rate is ∼1/2 the MC rate, it’s just the shape we want to compare here).
**Multiplicity**

**Multiplicity** - is telling us how many PMT pairs surpassed the threshold at least once in one event.

- The excess at low multiplicity for the real cosmic data can indicate the noise contribution or spill over from light still in the active volume when the trigger window activates.
- After applying a cut on Multiplicity bigger than 7 we observe a better agreement of data to MC samples.
**Comparison with the simulated cosmic muon data**

- **NOpening** - the number of times each PMT pair surpasses the trigger threshold per event. *

![Number of openings](image1)

![M > 7](image2)

The number of times each PMT pair *opens*. Number of openings after a cut on Multiplicity.

*The time interval we’re waiting to check if the new opening happened: 200 ns.*
**COMPARISON WITH THE SIMULATED COSMIC MUON DATA**

- **OpeningTime** - the time each PMT pair went above the trigger threshold per event.**

For all the variables: **Multiplicity**, **NOpening** and **OpeningTime** is clear that something is going on at low values.

**The larger Opening time, the later PMT pair crossed the threshold.**
Results (Binary Classification)

- For the training data sample* we used the ratio of cosmics to neutrinos as expected in the real data (SBN-doc-14145-v3):
  - ~ 1 $\nu$ interaction every 180 spills.
  - ~ 1 over 55 spills, is due to cosmic rays inside the beam spill time window.
- Updates to our methodology:
  - weighting of the loss function (allows more neutrinos in training),
  - running a more sophisticated training,
  - removal of empty cosmic events,
  - training on bigger statistics.

$$E = \frac{\#\nu \text{ tagged as } \nu \text{ in test sample}}{\#\nu \text{ in test sample}} = 91\%$$

$$P = \frac{\#\nu \text{ tagged as } \nu \text{ in test sample}}{\#\text{events in test sample}(\mu + \nu) \text{ tagged as } \nu} = 66\%$$

$$P_{\text{before CNN}} = \frac{\#\nu \text{ in training sample}}{\#\text{events in training sample}(\mu + \nu)} = 23\%$$

* Training and test sample consider triggered events only
CNN score = the probability of an event having particular label (ν or cosmic $\mu$).

ν purity after the CNN increased by a factor of 4.
**VISUALISATION OF THE IMAGES USED TO TRAIN OUR CNN**

Looks like long track with a systematic offset in the times on one wall with respect to another (presumably because the track is closer to one wall).

Predicted as neutrino event
runNo: 4642, subRunNo: 1, eventNo: 305, prediction: 0.263

Looks like it’s not through going muon (not so many adjacent PMTs in Z dir. are lit) and that has some hits at totally different times to the others (presumably coming from a second cosmic).

Predicted as cosmic background event
runNo: 4642, subRunNo: 1, eventNo: 63, prediction: 1.000
The classification is not biased by neutrino energy.

Neutrino selection efficiency with respect to the neutrino energy becomes high and flat for $E_\nu > 0.5$ GeV.
CNN PERFORMANCE W.R.T OUTGOING LEPTON ENERGY

- The classification is not biased by outgoing lepton energy.
- High and flat neutrino selection efficiency with respect to the outgoing lepton energy.
CNN PERFORMANCE W.R.T OUTGOING LEPTON ANGLE

- The classification is not biased by outgoing lepton angle.
- High and flat neutrino selection efficiency with respect to the outgoing lepton angle.
MC vs Real Data (Multiplicity Cut)

$\varepsilon_V = 83\%, \mathcal{P}_V = 62\%$ (MC)

No Multiplicity cut (DATA)

Test on the simulated data with $M > 7$

After cut: $M > 7$ (DATA)

Test on run 4642 with $M > 7$
ν SELECTION EFFICIENCIES (M > 7)

- High and flat neutrino selection efficiency for kinematic variables:
  - Neutrino energy,
  - Outgoing lepton energy,
  - Outgoing lepton angle.
INTRODUCING THREE TYPES OF EVENTS

Considering three types of events: neutrinos out of active volume (OAV), neutrinos in active volume (IAV) and cosmic muons.

Purity\textsubscript{\textit{v}_{IAV}} = 0.619774

Eff\textsubscript{\textit{v}_{IAV}} 0.812252

Neutrinos in AV
Neutrinos out of AV
Cosmics
Total
CONCLUSIONS AND NEXT STEPS

- New results (from the binary classification) show that we are able to reduce cosmic background from $\sim 77\%$ to $\sim 34\%$ whilst maintaining a neutrino interaction selection efficiency within the BNB window of $\sim 91\%$
- Applying multiplicity cut at $M > 7$ shows better agreement between data and simulation and improves the neutrino selection efficiency.
- Introducing the third type of events shows that:
  - we can trust the network selecting cosmic muons ($P_{\text{cosmic}} = 93\%$),
  - we dramatically reduce the cosmic background (from $\sim 77\%$ to $\sim 38\%$) and keep a high $\nu$ selection efficiency ($\sim 81\%$).
- Further separation of the relatively small remaining background can be done in higher level analyses (we can’t expect only PMTs to get us to 100% neutrino purity).
- Further steps:
  - Train the network with additional PMT information,
  - add CRT information,
  - add $e^-$ vs $\mu$ tag to the neutrino trees,
  - work alongside with relevant experts to implement improved simulations of the PMT responses.