Context Enriched Prong CNN performance studies in NOvA

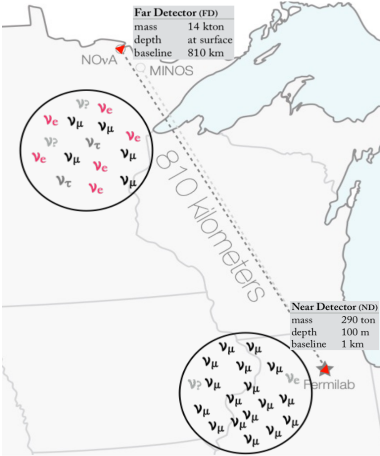
Akshay Chatla

New Perspectives 2021
16^{th}-19^{th} August 2021

August 19, 2021

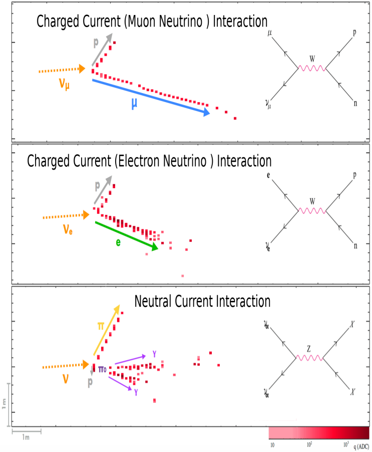
The NOvA Experiment

- NOvA is a long baseline neutrino experiment with a near detector at 1 km and a far detector at 810 km.
- The Physics goals:
  - study of neutrino oscillations ($\theta_{23}, \delta_{13}, \Delta m^2_{32}$)
  - neutrino cross-section measurements
  - exotic physics
- These goals need good neutrino interaction classification and neutrino energy measurement. Machine learning plays an important role in achieving them.



Event Classification At NOvA

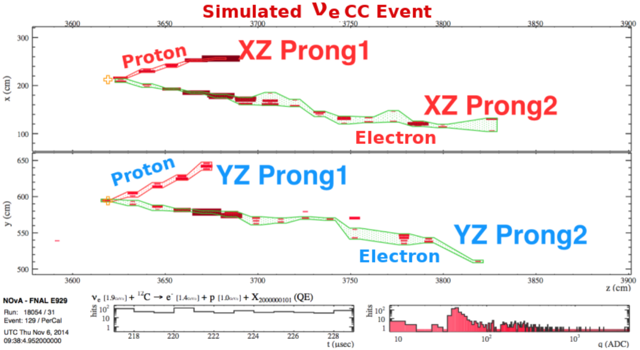
- NOvA was the first HEP experiment to use a Convolutional Neural Network (CNN) in a physics measurement to classify candidate neutrino interactions.
- NOvA uses Event CNN[1] which takes full event pixel map of both x and y views as input and the output is an event classifier.
- Event CNN can classify the events into 4 classes:
  1. ν_μ CC interaction
  2. ν_e CC interaction
  3. NC interaction
  4. Cosmics



[1] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa and P. Vahle, JINST **11** (2016) no.09, P09001 doi:10.1088/1748-0221/11/09/P09001

Particle classification

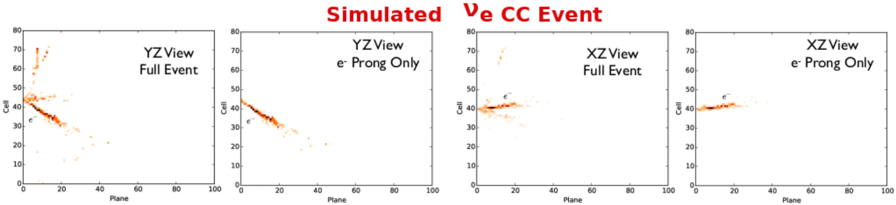
- While Event CNN can classify events, identification of final state particles of an event is needed to better our energy reconstruction and enable cross-section measurements of final states.
- Context enriched Prong CNN[2] is one of the networks being trained with a goal to identify all the final-state particles ($e^{\pm}$,p^{+}, $\mu, \pi^{\pm}, \gamma$) of a given neutrino event.



[2] F. Psihas, E. Niner, M. Groh, R. Murphy, A. Aurisano, A. Himmel, K. Lang, M. D. Messier, A. Radovic and A. Sousa, Phys. Rev. D **100**, no.7, 073005 (2019) doi:10.1103/PhysRevD.100.073005

Context enriched Prong CNN

- Prong CNN[2] uses a four-tower siamese-type MobileNetV2 architecture for including context information i.e it takes both event (Context) views and prong (Independent) views.
- Providing context to the network helps it to learn from physical quantities implying conservation laws as well as the relative topology of the particle with respect to the accompanying activity.



[2] F. Psihas, E. Niner, M. Groh, R. Murphy, A. Aurisano, A. Himmel, K. Lang, M. D. Messier, A. Radovic and A. Sousa, Phys. Rev. D **100**, no.7, 073005 (2019) doi:10.1103/PhysRevD.100.073005

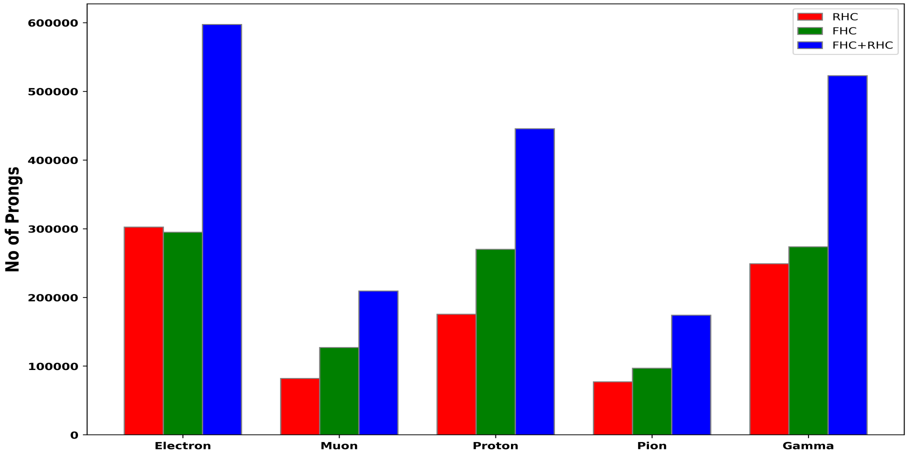
Motivation

- The previous iteration of Prong CNN used by NOvA trained on neutrino (FHC) and anti-neutrino(RHC) dataset separately. In this work, the network is trained on a combined (FRHC) neutrino and anti-neutrino dataset and compared with the separately trained networks.
- Following selection cuts are used on the training sample:

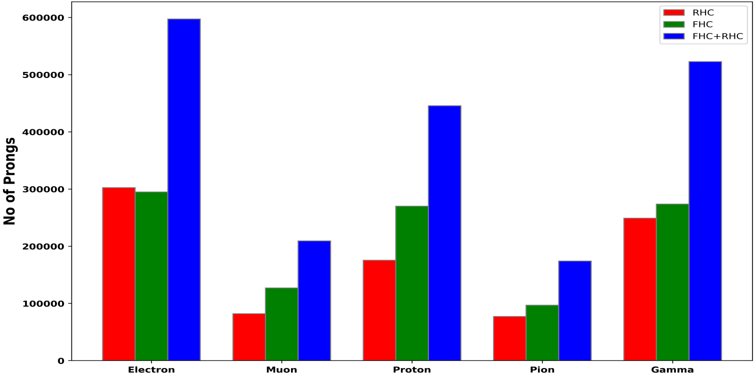
Selection cuts	Description
Containment Cut	Selects the prong and the event contained within the detector boundaries
Cosmic Veto	Removes cosmic events
Purity Cut	Realistic looking cluster with prongs $\mu^{\pm}, \gamma$: 50 %, $e^{\pm}$: 40 %, $\pi^{\pm}, p^{+}$: 35 %
Prong length	Cut prongs with prong length more than 5 m

Preliminary Training

- Before training the whole dataset, I took a part of FD simulated datasets for preliminary training.
  FHC had 1063406, RHC had 886712 prongs, Combined: 1950118



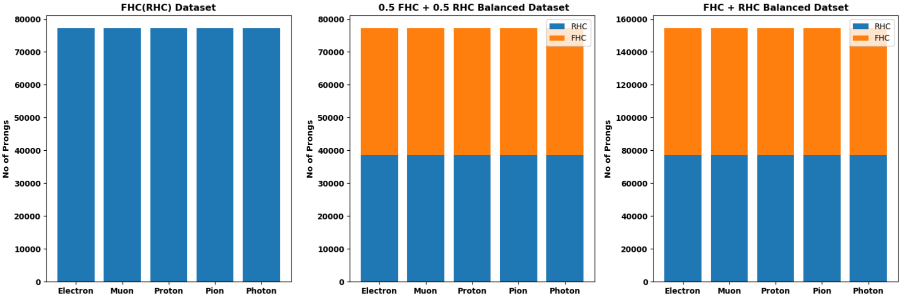
Preliminary Training



- This training dataset is unbalanced, one needs to train the network on a balanced dataset to make sure that network does not learn bias towards the class with more samples.
- Since, our training dataset is unbalanced, it is later balanced to contain an equal number of each type of particle.

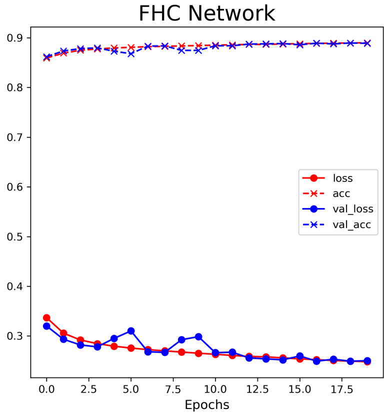
Preliminary Training

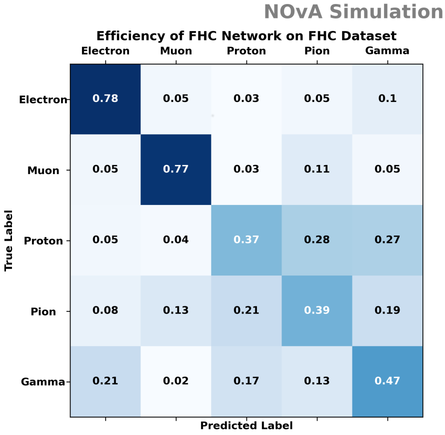
- Now, We have a choice of 4 kind of datasets we can train.
  1. A balanced neutrino dataset (FHC)
  2. A balanced anti-neutrino dataset (RHC)
  3. A half of $\nu + \overline{\nu}$ dataset (0.5 FHC + 0.5 RHC)
  4. A double $\nu + \overline{\nu}$ dataset (FHC + RHC)
- The balanced datasets look as follows:



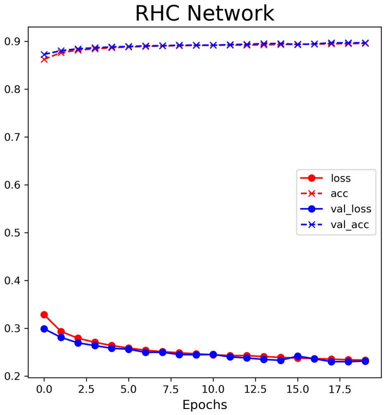
- For ease of comparison, I balanced the datasets to have same number of total events, the 4th dataset will have the double the data

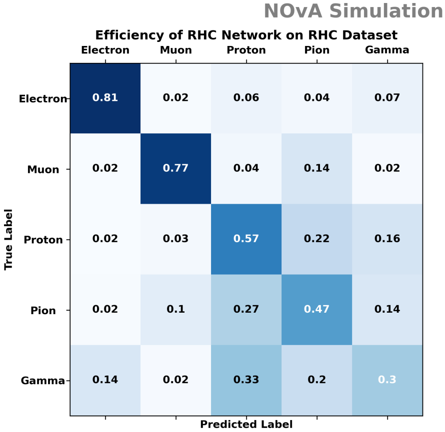
Only FHC Training/Evaluation



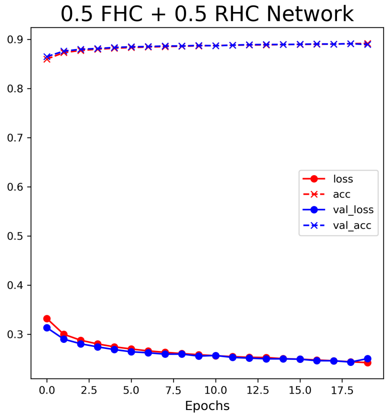


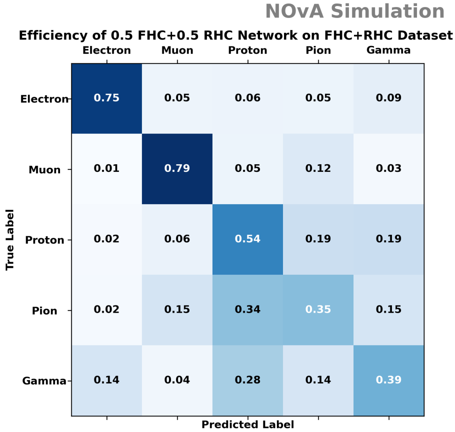
Only RHC Training/Evaluation



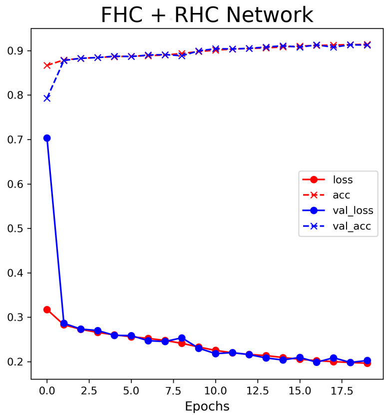


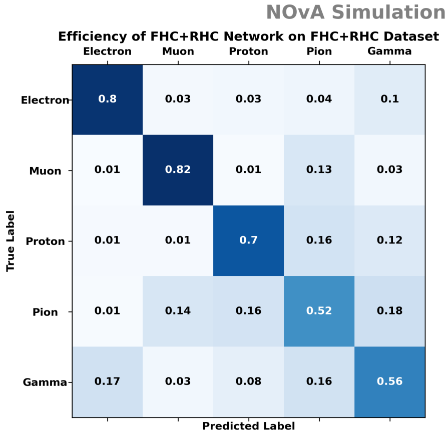
0.5 FHC + 0.5 RHC Training/Evaluation



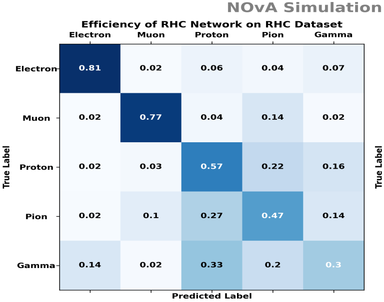


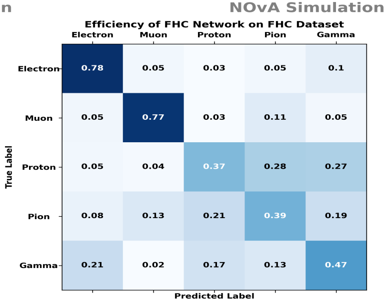
Double FHC + RHC Training/Evaluation

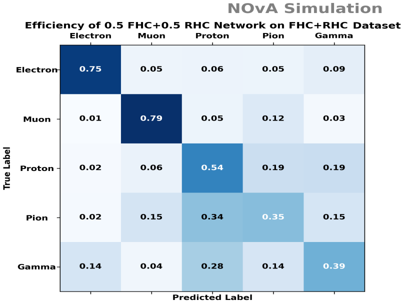




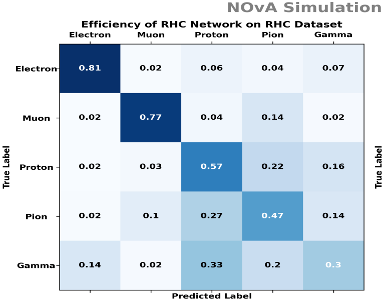
Network Comparison

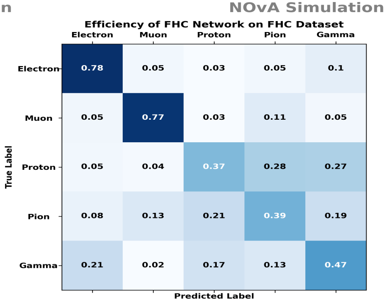


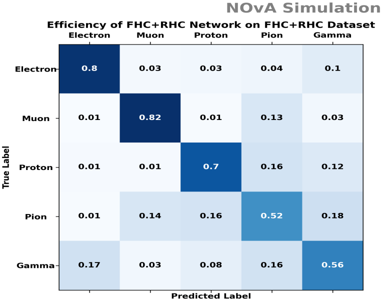




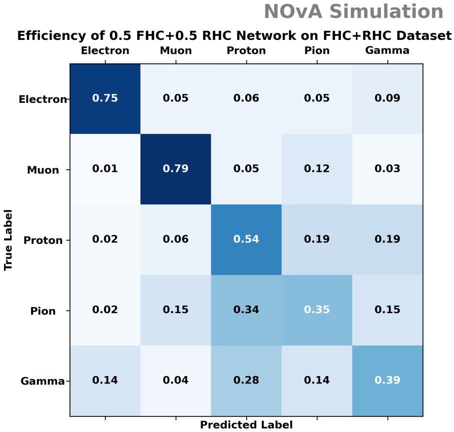
Network Comparison

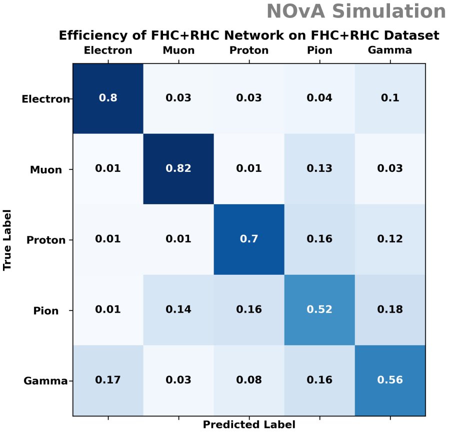




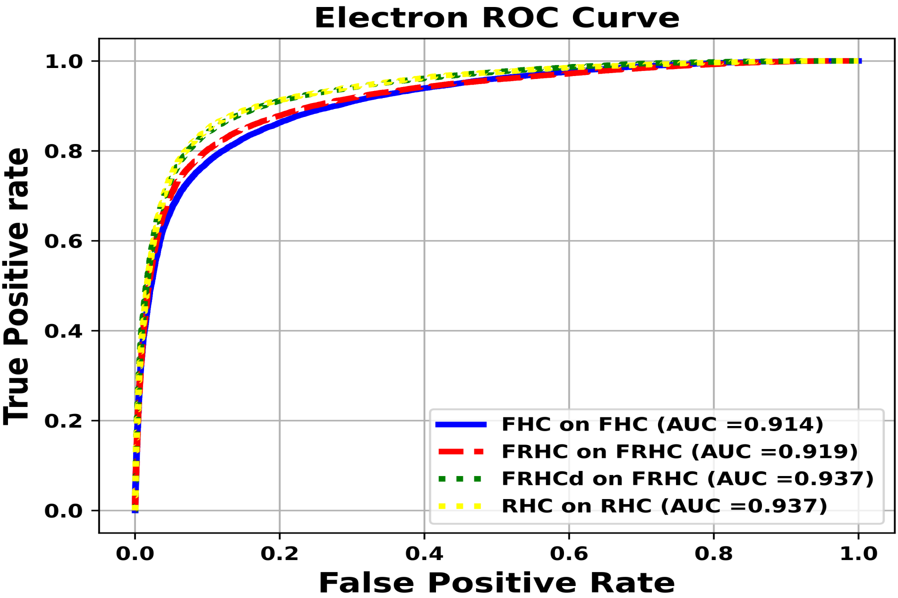


Network Comparison

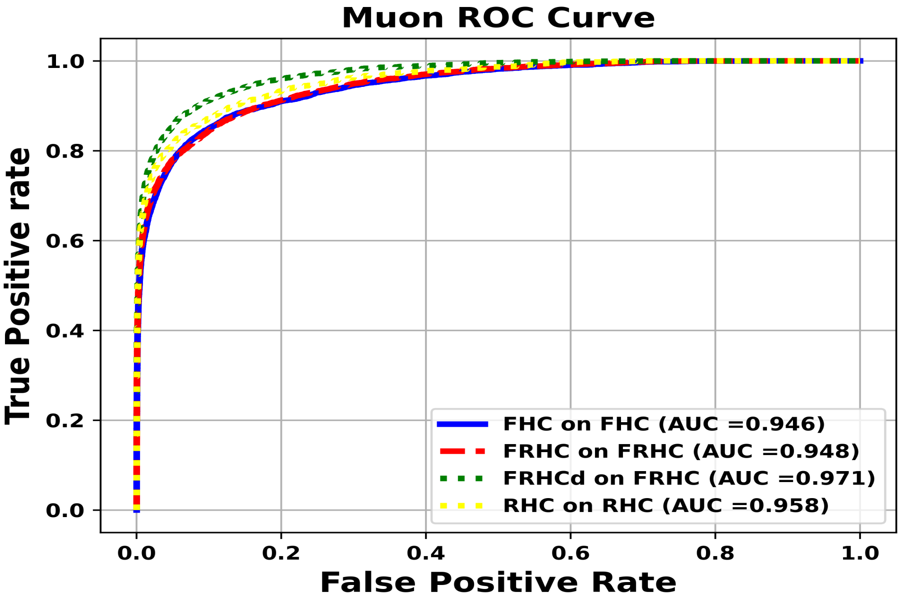




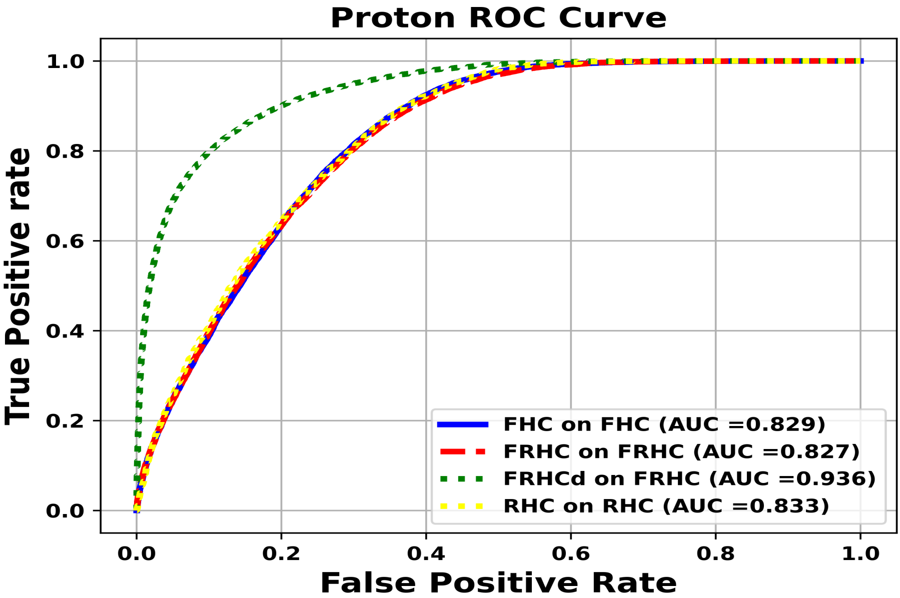
Electron ROC Curve



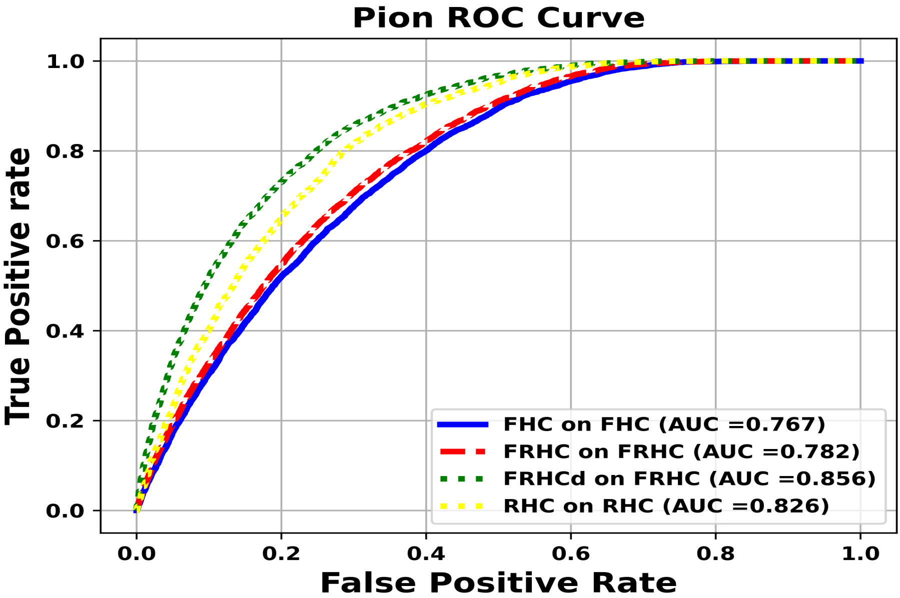
Muon ROC Curve



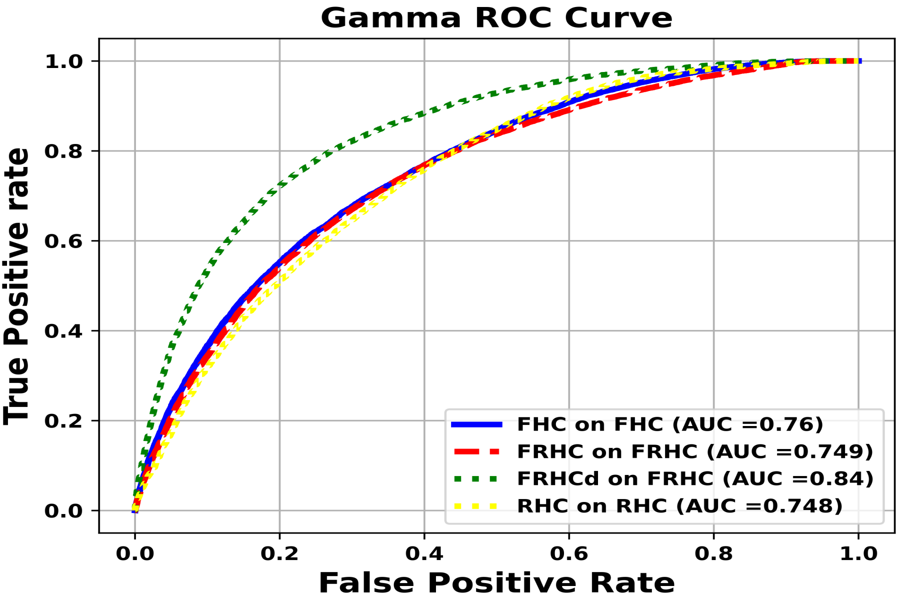
Proton ROC Curve



Pion ROC Curve



Gamma ROC Curve



Preliminary Conclusions

- Preliminary training networks show that double FHC + RHC network is comparatively better than RHC and FHC networks trained separately
- **Future:** Will try to find if the network performance can be improved by hyperparameter optimization.
- **Future:** Train with complete dataset and compare the three networks.
- **Future:** Check the effect of purity cuts on network performance.

THANK YOU

Backup Slides

Context enriched Prong CNN Architecture