ELECTRON AND PROTON CLASSIFICATION WITH AMS ECAL USING CONVOLUTIONAL VISION TRANSFORMERS AND DOMAIN ADAPTATION

B.Türk[1], M.B.Demirköz[1], E.Akbaş[2], Z.Weng[3], G.Karagöz[1]

[1]Middle East Technical University, Physics Department, IVMER, Ankara, Türkiye

[2]Middle East Technical University, Computer Engineering Department, Ankara, Türkiye

[3]CERN, Geneva, Switzerland

6th Inter-experiment Machine Learning Workshop in CERN, 29 January-2 February 2024

Abstract

Alpha Magnetic Spectrometer (AMS-02) is a precision high-energy cosmic-ray experiment on the ISS operating since 2011 and has collected more than 228 billion particles. Among them, positrons are important to understand the particle nature of dark matter. Separating the positrons from cosmic background protons is challenging above 1 TeV. Therefore, we use state-of-the-art convolutional and transformer models, CoAtNet and Convolutional Vision Transformer (CvT), that employ the shower signals from the ECAL to classify the electrons/positrons in the dominant cosmic proton background. We created sets of electrons, positrons, and protons events from the ISS data and Monte Carlo Simulation in the energy range between 0.2-2 TeV by applying various data quality cuts on reconstructed variables obtained from the subdetectors. Initially, since ECAL showers are not tunned in the AMS MC, our MC trained models show a lower proton rejection on the ISS data. To accommodate the difference between the training and test domain distributions, we implemented domain adaptation with the CoAtNet and CvT to mitigate this dataset bias/domain shift. We also trained domain adaptation with a set of well-reconstructed 1 electron charge ISS events without electron/proton labels at TeV energy order as the target dataset. We evaluated the models between 1-2 TeV energy using ISS and MC events with the proton rejection vs. electron efficiency and proton rejection vs. energy at near 90% electron efficiency plots. We performed experiments using various training and validation dataset combinations and other hyperparameters with the CvT and CoAtNet. Among them, the best models are obtained with the 1-2TeV MC events as training data and half of the labeled 1-2 TeV ISS events as validation data. Using domain adaptation with the CoAtNet, we obtained a maximum proton rejection at 88% electron efficiency on the ISS data. We also rejected all of the MC protons at higher than 99.8% electron efficiency with both CvT and CoAtNet. At 90% electron efficiency, the proton rejection power of the CvT and CoAtNet is 5 and 7 times higher than the proton rejection power of the AMS's Boosted Decision Tree and ECAL Likelihood Estimator for MC events in the 1-2 TeV range.

Introduction

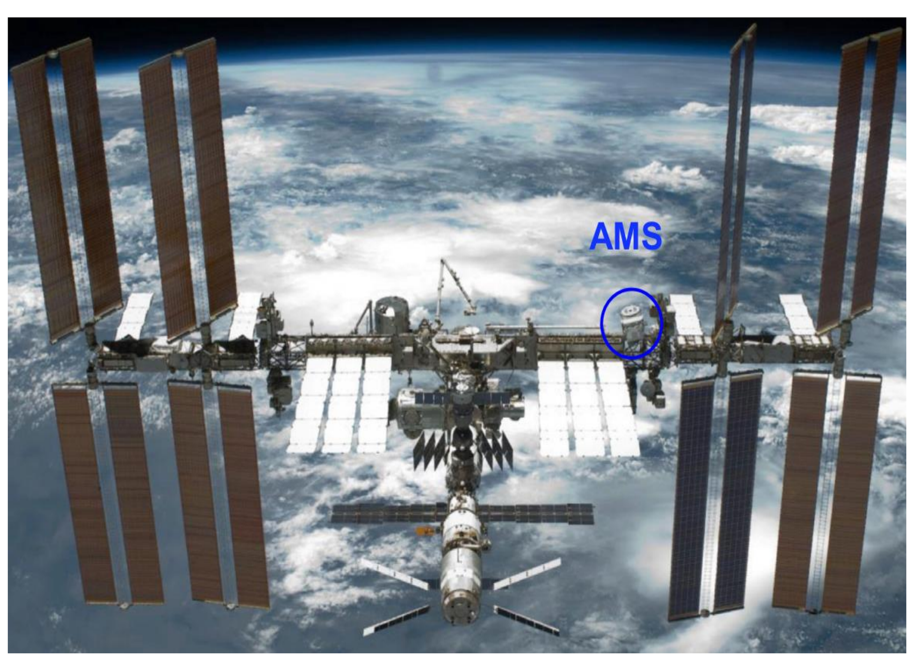

Figure 1: AMS-02 on the International Space Station [1].

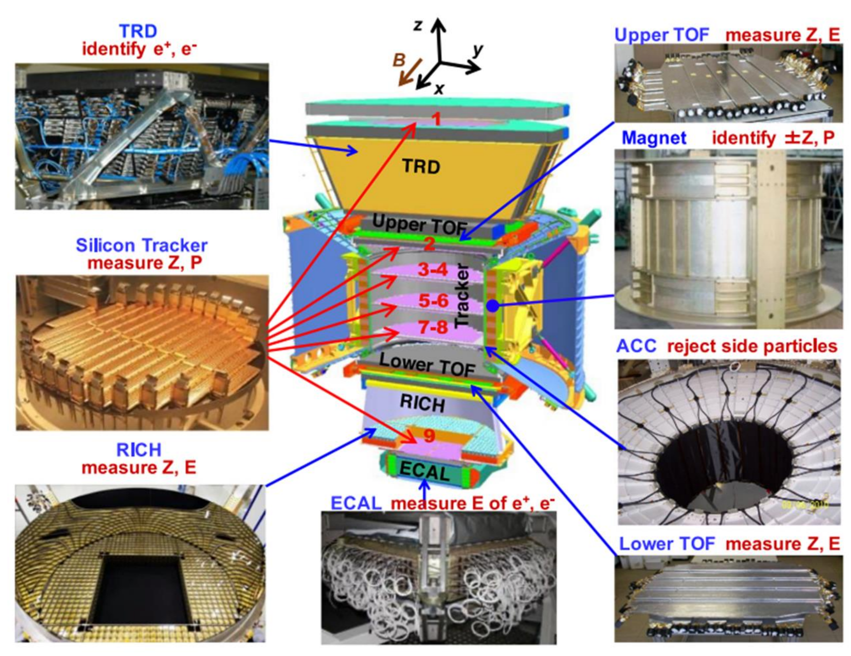

Figure 2: AMS detector with the subdetectors [1].

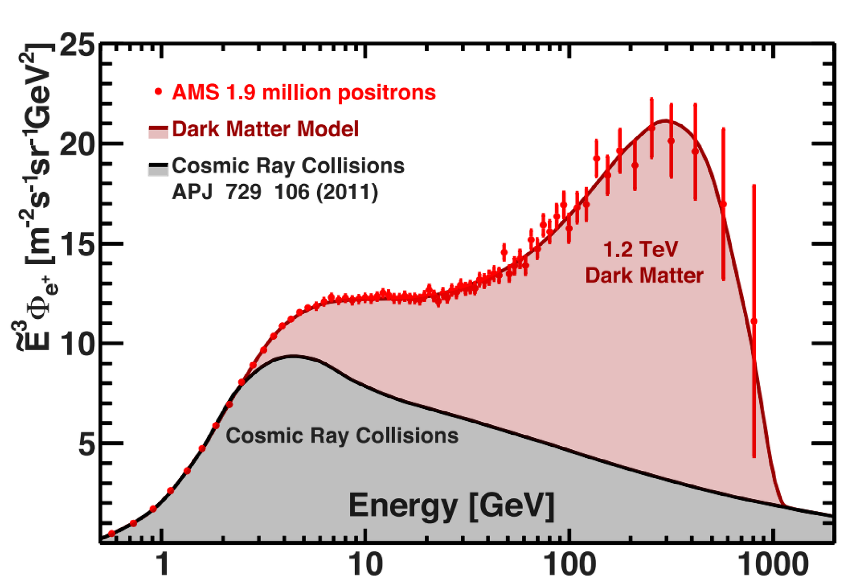

Figure 3: Positron Spectrum measured by AMS-02 [1].

- AMS observed more than 228 billion cosmic ray events.
- AMS measurements show an excess of positron flux above 25 GeV higher than what is expected from cosmic ray collisions [1].
- To confine the positron models better, we need high positron statistic with high proton rejection above 1 TeV
- Positron separation from background protons are challenging above 1 TeV;

1. There are more than 10^4 protons per positron
2. Proton showers consist of electromagnetic showers because π^0 decays to 2 photons.
3. TRD's proton rejection power decreases
4. Recons. Energy/Rigidity ratio deviates more from 1 for the e+ and e- due to the Tracker

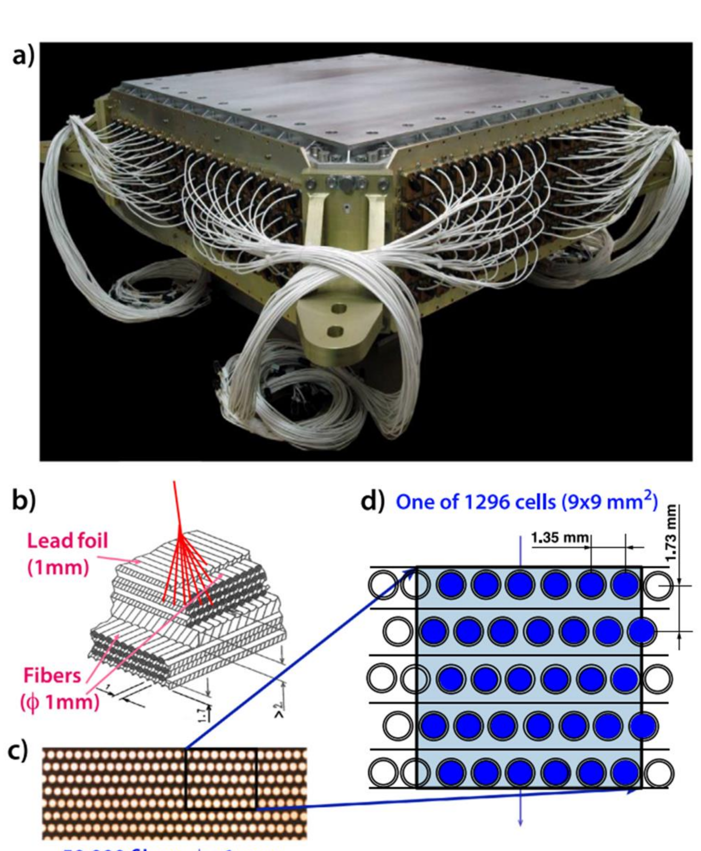

Figure 4: (a) AMS Electromagnetic Calorimeter. It is 17 X_0 (b) Small section of a super layer of the ECAL. (c) Scintillation Fibers. (d) A cell which is collected by the one of anode of the PMT [1].

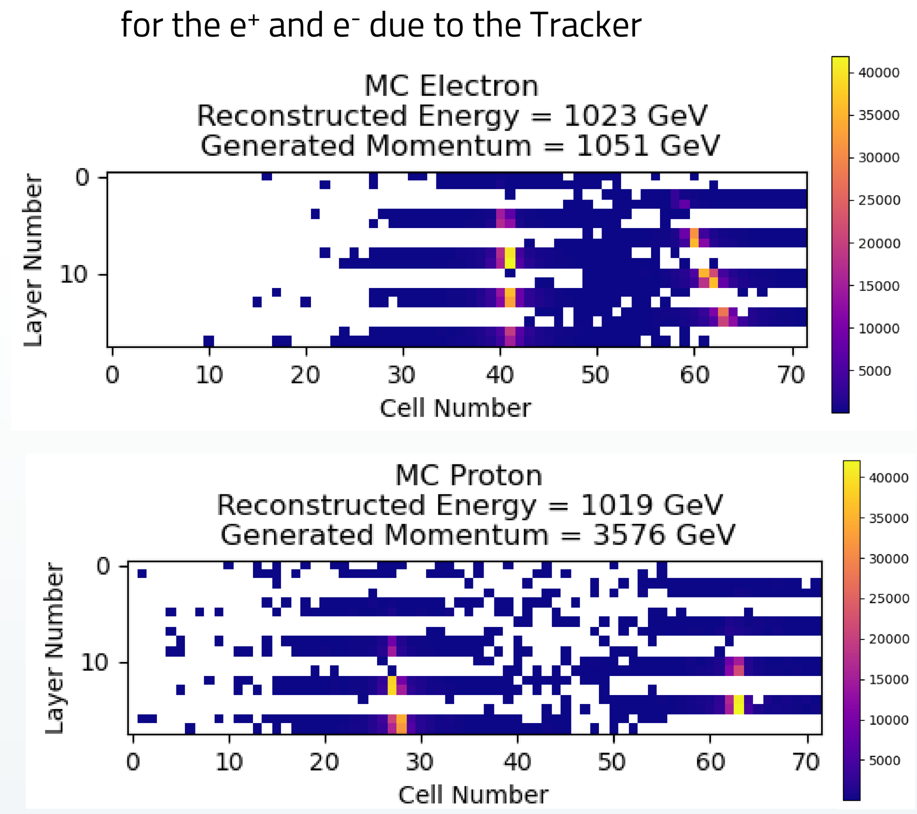

Figure 5: ECAL shower signals. When the electrons and protons pass through the electromagnetic calorimeter, many secondary particles are created. The particles deposits energies which are recorded by the photomultiplier tubes.

Method

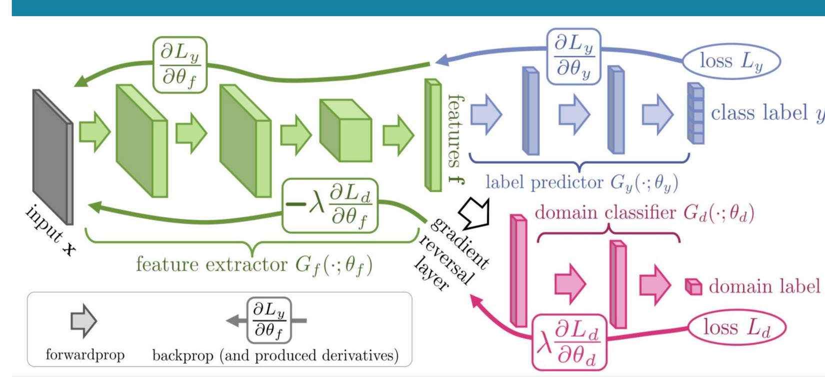

Figure 6: Architecture of Domain Adaptation. It consists of 3 networks, Feature Extractor, Label Classifier, and Domain Classifier [2].

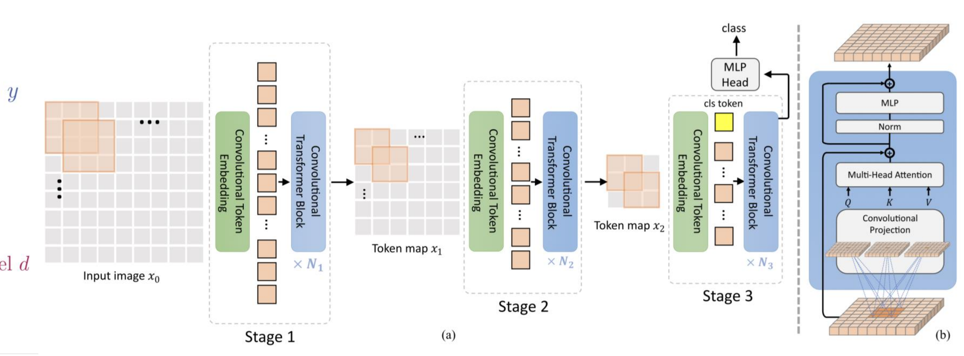

Figure 7: (a) Architecture of Convolutional Vision Transformers. (b) Convolutional Transformer Block [3].

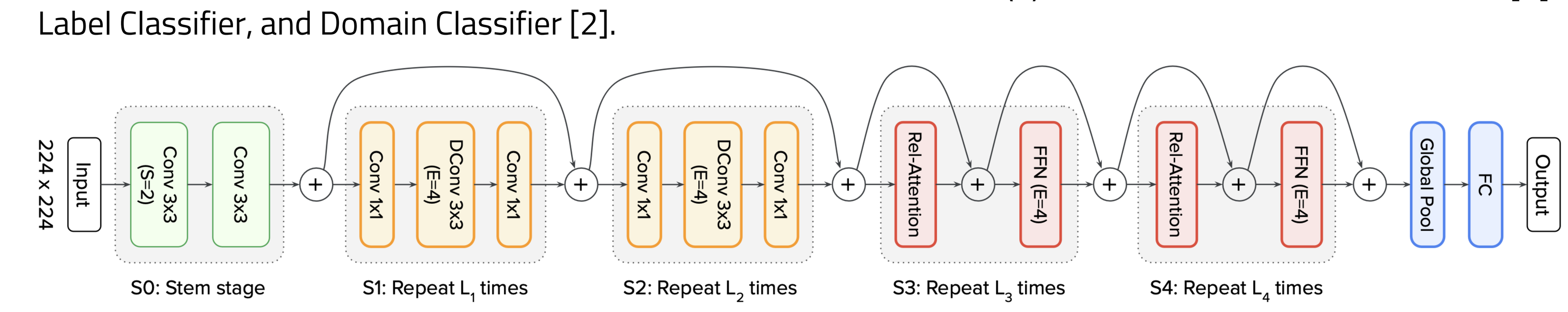

Figure 7: Architecture of CoAtNet [4].

- Early Stopping to prevent the overfitting
- Weighted Cross Entropy Loss Function for the class imbalance
- Adam Optimization Algorithm [5]
- Domain Adaptation to decrease the domain shift between MC and ISS Dataset
- Created training, validation, test and target sets by processing raw data and applying cuts to obtain good reconstructed events with high purity.

	MC Lower Energy Set	MC Higher Energy Set	ISS Validation Dataset	ISS Test Dataset	ISS Target Set (Well reconstructed 1e- ISS Data)
Number of Electrons	3.5×10^6	1.2×10^6	10	11	Labels are not necessary
Number of Protons	2×10^6	2.6×10^6	828	696	Label are not necessary
Total Number of Samples	5.5×10^6	3.8×10^6	838	707	2×10^5
Reconstructed Energy Range	$0.2-1\,TeV$	$1-2\,TeV$	$0.95-2\,TeV$	$1-2\,TeV$	$0.2-2\,TeV$

Table 1: Details of ISS Datasets and MC sets

Results & Conclusion

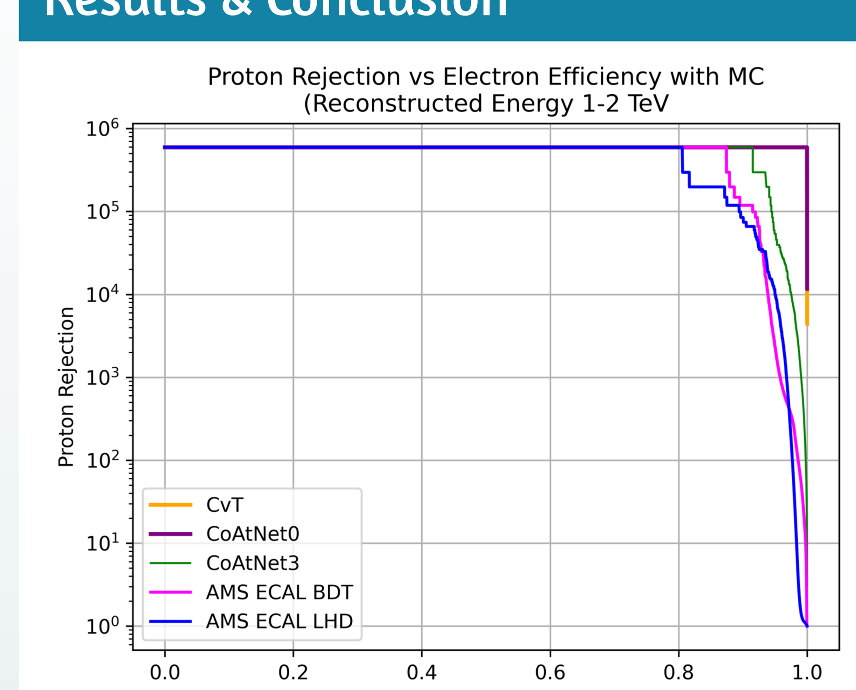

Figure 8: Proton Rejection vs. Electron Efficiency with MC Test Set(1-2 TeV).

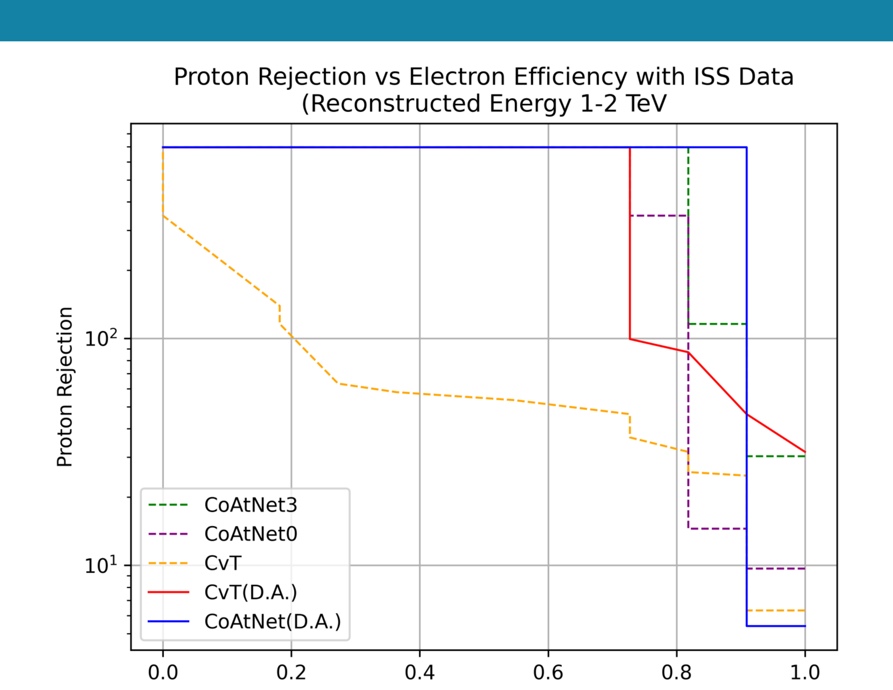

Figure 9: Proton Rejection vs. Electron Efficiency with ISS Test Dataset(1-2 TeV).

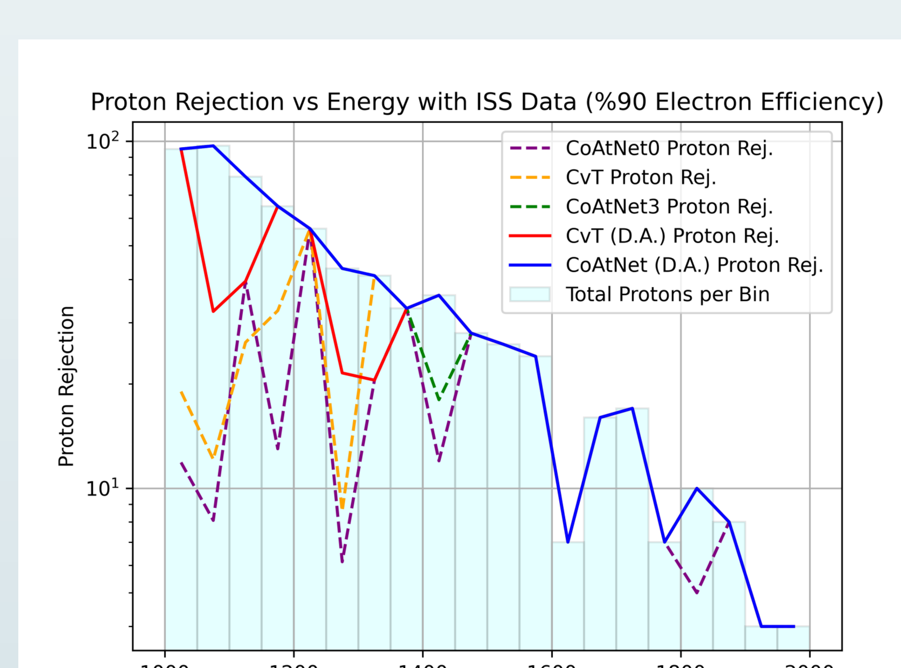

Figure 10: Proton Rejection vs. Energy at 90% Electron Efficiency with ISS Test Dataset.

$$Proton\ Rejection = \frac{Total\ Number\ of\ Protons}{Total\ Number\ of\ Misclassified\ Protons}$$

- CvT, CoAtNet0, and CoAtNet3 are trained with MC Training Set (0.2-1 TeV).
- MC Validation(1-2 TeV) set for CvT and CoAtNet0.
- ISS Validation set for CoAtNet3
- CoAtnet3 misclassifies 6 ISS Data protons at around 90% electron efficiency.
- Better Results are obtained for CvT(D.A.) and CoAtNet(D.A.) by using MC High Energy .Set.
- CoAtNet with Domain Adaptation has only 1 misclassified proton on the ISS Test Dataset.

References

1. M. Aguilar, L.A. Cavasonza, G. Ambrosi et al. Physics Reports 894 (2021) 1
2. Ganin, Yaroslav, and Victor Lempitsky. "Unsupervised domain adaptation by backpropagation." International conference on machine learning. PMLR, 2015.
3. Wu, Haiping, et al. "Cvt: Introducing convolutions to vision transformers." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.
4. Dai, Zihang, et al. "Coatnet: Marrying convolution and attention for all data sizes." Advances in neural information processing systems 34 (2021): 3965-3977.
5. Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980*(2014).

Acknowledgements

This project is supported by The Scientific and Technological Research Council of Türkiye (TÜBİTAK) under Program Name 1001 - The Scientific and Technological Research Projects Funding Program, Project No. 123F090