Application of Machine learning in Muon Tomography

Raman Sehgal[1],* Aniket Shinde[2], L. M. Pant[1], and B. K. Nayak[1]
[1] *Nuclear Physics Division, Bhabha Atomic Research Centre, Mumbai - 400085, INDIA and*
[2] *Vivekanand Education Society's Institute of Technology, Mumbai, INDIA*

Introduction

Muon tomography very promising technique to scan the cargo containers and to detect the presence of high Z materials. Existing algorithms follow iterative approach to do the image reconstruction and to discriminate between different materials. The class of problems relates to a very well known classification problem in Machine learning[1]. Binary classification technique of machine learning can be used to get a clear reconstructed image by subtracting the background from the data, on the other hand multiclass classification can be used to discriminate and identify different materials under test. We will present the preliminary results of application of machine learning algorithm to do the image reconstruction and material identification on simulated data generated using Geant4 [2].

1. Machine Learning

Solving a problem using computers needs the problem logic to programmed so that it can be executed on the computational processor. Machine Learning is the substream of computer science that gives these computational devices the capability to solve the desired problem without being explicitly programmed to do that. It works on data driven model and try to solve the problem by learning the model from the exposure of existing data. The more the data, the more accurate model it can build. Once the model is ready then it can be used to solve the problem at hand. Machine learning algorithms are broadly classified in following two categories 1) Supervised Learning and 2) Unsupervised learning. The difference between these algorithms is the way in which they build the model from the data.

Supervised algorithms needs labeled (annotated) data to build the underlying statistical model. On the other hand, the unsupervised algorithm does not needs labeled data and try to find the hidden pattern itself.

2. Machine Learning in Muon Tomography

Our idea of using machine learning in muon tomography is to do the clear image reconstruction as well as to do the material discrimination and identification. The image reconstructed using Point of Closest Approach algorithm (PoCA) [3] contains a lot of false positive which acts as background and result in smeared image. Using binary classification technique of machine learning the signal and background can be segregated. The subtraction of identified background results in clear reconstructed image. Similarly material identification can be done by using multiclass classification technique, where each class corresponds to different material. The total number of classes in multiclass classification is equal to number of material under which classification needs be done plus one more class for background.

3. Data Sets

The data sets are generated using Geant4 simulation. Here we are going to use supervised machine learning algorithms. The training datasets are generated by placing three cubical block of dimension($20cm \times 20cm \times 20cm$) and of Al, Fe and Pb material. These are kept at (0cm, 30cm, 0cm), (-30cm, -30cm, 0cm) and (30cm, -30cm, 0cm) locations respectively as shown in figure 1. From each of the incoming and outgoing tracks 4 features are calculated. These are angular deviation along

*Electronic address: rsehgal@barc.gov.in

X direction ($\Delta\theta_x$), angular deviation along Y direction ($\Delta\theta_y$), linear deviation in X direction (Δ_x), linear deviation and Y direction (Δ_y). In the simulation we can also get the momentum(ρ) of muon, so it is also stored as one of the features ie. a vector of total 5 feature is available per muon trajectory. To make the data labeled we have also stored the class to which the corresponding data point belongs. Test dataset is generated by placing the scatterers at different position from training dataset, and storing the same features as for training dataset. These dataset are stored in ROOT [4] tree format.

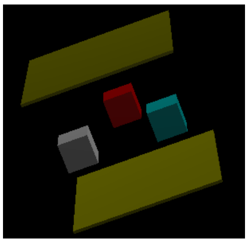

FIG. 1: Training Simulation setup

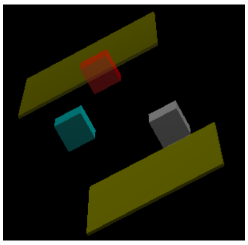

FIG. 2: Testing Simulation setup

4. Results

The results are presented for three cubical block of dimension($20cm \times 20cm \times 20cm$) of Al, Fe and Pb material. These are kept at (0cm, 30cm, -30cm), (-30cm, -30cm, 10cm) and (30cm, 0cm, -10cm) respectively. Various machine learning algorithm are compared for the accuracy of classification of PoCA points as shown in table I. Figure 3 shows the raw PoCA points, where there are a lot of false positives. Figure 4 show the filtered output after removal of points which are classified as background using Random Forest algorithm. The color of PoCA point in figure 4 corresponds to the material class to which the PoCA point belongs. Following color conventions are used, Red for Al, Cyan for Pb and White for Fe. Here we can clearly distinguish between low Z (Al) and high Z (Pb) material. For medium Z material (Fe) a considerable proportion of PoCA points are predicted as PoCA points belongs to either Al or Pb. This can be resolved by collecting more data. Training of the algorithm also depends on the training data available for each class. The more the data we collect for each of the training class the better prediction we can do.

Algorithm Name	Classifiation Acc.(%)
Decision Tree	87.41
Random Forest	87.43
K-Nearest Neighbours	87.61
Linear Discriminant Analysis	83.80

TABLE I: Classification accuracy of different algorithms on simulated data.

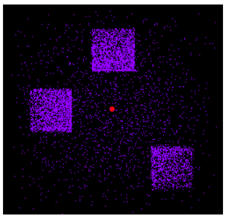

FIG. 3: Raw PoCA points from Simulated data

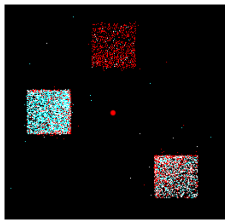

FIG. 4: Background rejection using Random Forest

Conclusion

In this paper we have show the application of Machine Learning algorithm in muon tomography. Classification performance of various algorithms are compared, and results of application of Random Forest algorithm for background rejection and material identification are shown. In future we will try to apply these algorithm on the experimental data.

References

[1] https://arxiv.org/pdf/1808.02342.pdf
[2] S. Agostinelli et al., *GEANT4, NIM-A*, **506** (2003) 250.
[3] Raman Sehgal et al. *Proc. of the DAE Symp.on Nucl. phys. Vol 63 (2018)*
[4] Rene Brun and Fons Rademakers, *ROOT* , *NIM-A* **389** (1997) 81-86