

Neural Networks for the Reconstruction and Separation of High Energy Particles in a Preshower Calorimeter

Juan Pavez¹(✉), Hayk Hakobyan², Carlos Valle¹, William Brooks²,
Sergey Kuleshov², and Héctor Allende¹

¹ Departamento de Informática, Universidad Técnica Federico Santa María,
CP 110-V, Valparaíso, Chile
juan.pavezs@alumnos.usm.cl

² Departamento de Física, Universidad Técnica Federico Santa María,
CP 110-V, Valparaíso, Chile

Abstract. Particle detectors have important applications in fields such as high energy physics and nuclear medicine. For instance, they are used in huge particles accelerators to study the elementary constituents of matter. The analysis of the data produced by these detectors requires powerful statistical and computational methods, and machine learning has become a key tool for that. We propose a reconstruction algorithm for a preshower detector. The reconstruction algorithm is in charge of identifying and classifying the particles spotted by the detector. More importantly, we propose to use a machine learning algorithm to solve the problem of particle identification in difficult cases for which the reconstruction algorithm fails. We show that our reconstruction algorithm together with the machine learning rejection method are able to identify most of the incident particles. Moreover, we found that machine learning methods greatly outperform cut based techniques that are commonly used in high energy physics.

Keywords: Machine learning · Neural networks
Reconstruction algorithm · Computational physics

1 Introduction

Event reconstruction is an important task in pattern recognition. The task is inherently statistical and consist on inferring an event given the data that it produced. It has applications in fields such as *high energy physics*, *nuclear medicine* and *astrophysics*.

In particular, in *high energy physics* event reconstruction is used to analyze the particles produced in particle accelerators. These accelerators collide particles at high velocity, producing a huge amount of secondary particles. The observation and study of the particles produced may help to understand the structure of matter and its fundamental properties. The particles produced at

collisions must be measured by detectors attached to the accelerators. These detectors quantify different properties of the particles, such as energy, momentum or interaction time.

To infer the particles that were produced, an event reconstruction algorithm is used. The reconstruction algorithm must process the information produced by the detector in order to identify the particles that originated the data. To do that, the algorithm must be able to distinguish between signal and background, cluster together data from each particle, separate overlapping particles, classify data from different particles, among others. Such tasks require powerful computational and statistical methods, and machine learning has become an important tool to solve many of these problems [1–3].

In this work we propose a complete reconstruction algorithm for a detector under construction at the Scientific and Technological Center of Valparaíso (CCTVal). More importantly, we propose the use of machine learning, particularly neural networks, to solve the problem of identifying overlapping particles. The detector and reconstruction algorithm can be used to identify close particles produced in collisions of electron ion colliders and also have applications in nuclear medicine.

2 Preshower Detector

The identification of neutral pions is an important problem in high energy physics. The task is difficult because pions decay into two photons with a very small opening angle. Consequently, the two photons arrive very close to each other to the detector, which makes very hard to distinguish between the decaying neutral pion and a high energy photon. The impossibility to identify neutral pions might affect the whole reconstruction process, since those particles can contribute important clues to the understanding of the underlying process.

The main limitation of commonly used detectors is the spatial resolution. To solve that problem a preshower calorimeter detector is proposed in [4]. This kind of detectors work by generating electromagnetic particle showers when an incident particle is detected. The showers are measured using a readout mechanism that quantifies the energy produced by the showers. For that reason, and given that the detector is used in front of the main detector, it is called preshower detector. The preshower detector is composed by a matrix of scintillating crystals, which are the ones that produce the electromagnetic showers. The signal of the showers is conducted by a set of optical fibers to the readout mechanism located on the sides of the front face of the crystal matrix. The readout system is composed by two vectors of measuring cells, producing two vectors of energy counts. The smaller transversal size of the crystals used in the preshower allows a better spatial resolution, helping with the identification of particles produced by neutral pions.

3 Reconstruction Algorithm

The outputs of the readout are two vectors of photoelectron counts, from where the energy can be computed. Each vector has 25 values, corresponding to each measuring cell. The vectors can be considered X-Y projections of the energy deposited on the crystal matrix. Using that information the reconstruction algorithm must be able to reconstruct the energy and position of the incident particles. The reconstruction process is composed by a series of steps which are resumed in Fig. 1. Each one of the steps are explained below:

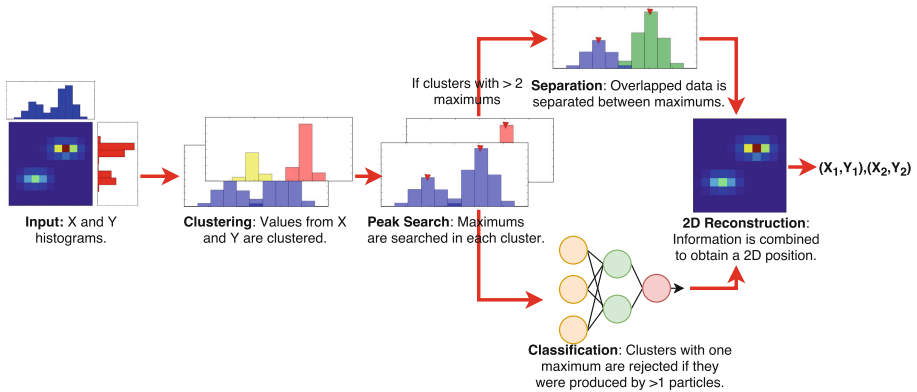


Fig. 1. Diagram of the full reconstruction process.

Clustering algorithm: The process starts with a clustering algorithm that must identify and separate each one of the incident particles. This is done by clustering together the values on the cells and assuming that each cluster corresponds to either one incident particle or two or more overlapping particles. The study of clustering algorithms is extensive and there are plenty of options to use in this step, including methods specially designed for clustering in calorimeters [5] or general pattern recognition methods [6]. We decided to use a simple, but efficient method called *topological clustering*. The method starts with a single cell, commonly the one with the maximum energy. Then, it iteratively adds the neighbors of the cells already included. Cells are only added if they are above an acceptance threshold which depends on the expected noise. Special care must be put in the process of combining clusters constructed for each one of the axes, since ambiguities may show up when two or more hits are spotted. To solve that problem we observe that clusters with similar energies in different axes were probably produced by the same particle. Using that assumption, we solved the ambiguities problem by considering all possible combinations and then choosing the ones with almost equal energy deposition.

Peak finding: In a second step, a peak finding algorithm is used to find the maximums in each one of the clusters. More than one maximums might indicate

the presence of overlapping particles. We used the peak finding method proposed in [7]. The algorithm assumes that the peaks can be approximated by a normal function and that the background is piecewise linear. Then, peaks are identified using the second differences of the values, since the background is removed at the second derivative of the function.

Separation algorithm: If two or more maximums are identified in a cluster, a separation algorithm is used to set apart the overlapping particles. A separation algorithm for calorimeters is proposed in [8]. The algorithm is based on the lateral response function of the electromagnetic shower. That function relates the energy deposited in each cell with the distance from that cell to the incident particle position. Then, the function is used in an iterative algorithm that distributes the energy of the overlapping cells. The distribution is done proportionally to the normalized value of the lateral response function multiplied by the total energy deposited in the cell. The algorithm proceeds by alternating between re estimating the particle position and distributing the energy of each cell. We estimated the lateral response function for the preshower using simulations and we used the described algorithm to separate the overlapping clusters.

Position reconstruction: After identifying the energy corresponding to each particle, the position of the particle must be reconstructed. Using the location of each measuring device and the deposited energy, the reconstruction algorithm must estimate the position of each one of the detected particles. For that, we used the center of gravity of the cluster, but with logarithmic weights [9]. The logarithm accounts for the exponential decay of the electromagnetic showers.

Classification algorithm: Finally, a classification algorithm is used to reject cases in which only one maximum is observed for two incident particles. That is because the separation algorithm can only separate overlapping showers when two or more maximums are detected, however, two particles can be observed with one maximum if they are too close to each other.

4 Neural Networks for Particle Separation

In cases for which only one maximum is observed, it is hard to separate the overlapping showers. Nevertheless, it is possible to reject overlapping showers detected as single particles. In [8] the authors propose to use a cut in the second central moment of the shower, or dispersion, defined as

$$D_x = \frac{\sum E_i x_i^2}{\sum E_i} - \left(\frac{\sum E_i x_i}{\sum E_i} \right)^2. \quad (1)$$

It is observed, however, that the cut efficiency is dependent on the particle incident position respect to the center of the cell and on the energy of the particle. To solve the first issue, a parabolic cut on D_x is proposed, by defining the values

$$D_x^{corr} = D_x - D_x^{min}, \quad (2)$$

$$D_x^{min} = (\bar{x} - x_R)(\bar{x} - x_L), \quad (3)$$

where \bar{x} is the first moment of the cluster and x_R, x_L are the right and left edges of the central cell of the cluster. A linear cut in D_x^{corr} is equivalent to a parabolic cut in the distribution of D_x versus \bar{x} . While the method is useful to solve the dependence on the incident position, it is noticed that the rejection efficiency is energy dependent even when using D_x^{corr} .

More advanced multivariate classification techniques can be used at this stage. In [10,11], the authors propose to use machine learning methods in the task of particles discrimination in calorimeters. In both papers the authors derive a set of features from the showers and then use those features to train a multivariate classification method.

For the preshower it is interesting to study the use of multivariate classification methods. Multivariate methods can: (1) Provide more complicated nonlinear cuts that might improve the rejection performance when compared to simple linear cuts. (2) Obtain a rejection efficiency that is invariant with the incident energy. The latter can be done by using the total energy as input feature. That allows the classifier to varies smoothly for different incident energies, avoiding the energy dependence of the dispersion cut. Similar ideas have been presented in [12, 13]. Based on that, we propose to use a classifier trained on a set of features extracted from the cluster values. In particular, we propose to use a *multilayer perceptron* trained on 7 features extracted from the X and Y projections of the shower: the position mean and variance, the position skewness, the normalized height of the shower defined as $E_{max}/\sum E_i$ and the corrected dispersion. Each one of the features is computed for the X and Y projections obtaining a total of 14 features. Moreover, the total energy of the incident particle is included in the feature vector.

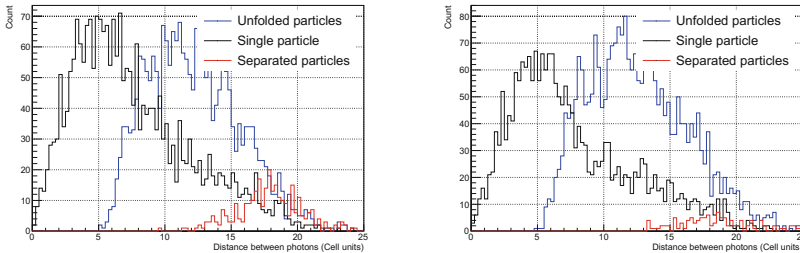


Fig. 2. Number of clusters identified as single particles, two overlapping or two separated particles given the distance between the particles. The experiments were performed for: 10 GeV + 1 GeV (left), 10 GeV + 10 GeV (right).

5 Experiments

To study the performance of the preshower detector a computational simulation of the detector was built using *Geant4* [14]. The experiments consisted on throwing photons on the front side of the preshower detector. Then, all the physics that occur in the crystals matrix and readout system were simulated. Finally, the

reconstruction algorithm was applied in the simulated measures of the readout cells. Simulations were carefully calibrated in order to obtain measures that are similar to the real experimental setup. The use of simulations is needed since is the only way to know exactly the event that originated the measures.

We evaluated the performance of the reconstruction algorithm in simulated pairs of photons with uniformly distributed incident positions and with energies 10 GeV + 1 GeV and 10 GeV + 10 GeV. In Fig. 2 we show the number of particles reconstructed given the distance between both incident particles. The options are: (1) A single cluster. (2) Overlapping clusters. (3) Two separated clusters. Since all particles were simulated in pairs, the detector obtains a correct result if it identifies two overlapped or separated clusters. It is noticed that those cases are mainly observed for distances ≥ 28 mm (or equivalently 7 cell units of size 4 mm). On the other hand, for distances ≤ 20 mm most of the particles are misidentified as a single cluster.

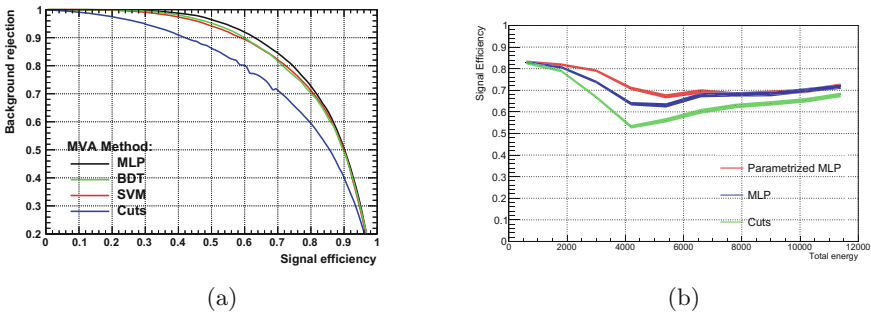


Fig. 3. (a) ROC curves for each one of the tested methods. (b) Signal efficiency for: (a) A neural network with total energy as input feature (parametrized), (b) A neural network without using the energy, (c) The cuts based method.

Next, we studied the rejection algorithm for particles with distances ≤ 20 mm, where the reconstruction algorithm fails to identify two particles. For that, we simulated single particles and pairs of particles. The algorithms were optimized to identify the cases of two particles. To evaluate the performance of the multivariate methods we will compare various machine learning algorithms (including *support vector machines*, *boosted decision trees* and *multilayer perceptron*) to the method based on cuts on the dispersion variable that has been previously used in [8].

For each one of the machine learning methods the features were normalized to the range $[-1, 1]$. We included the total energy for each classifier. For the *support vector machine* we used a *Gaussian* kernel with regularization $C = 1.0$. In the case of *boosted decision trees*, we used *AdaBoost* with 200 trees. We also used a maximum depth of 3 and learning rate of 0.5. For the *multilayer perceptron* we used an architecture of two hidden layers of size 15 and 5. We used *tanh* activations for the hidden layers and *sigmoid* activation for the output layer.

To train we used *stochastic gradient descent* with a learning rate of 0.02, ℓ_2 regularization and 600 epochs. The cuts were automatically selected in order to maximize the background rejection and signal efficiency, (r_B, e_S) . For that, the distribution of (r_B, e_S) distribution estimated using *monte carlo* sampling. Then, the best cut value for each one of the features was selected. It was observed that the method works poorly with many features, because of that, we used only the dispersion and corrected dispersion. All methods were implemented using TMVA [15].

The ROC curves for each one of the methods are shown in Fig. 3a. It can be seen that the best results are obtained by the *multilayer perceptron*.

Next, we studied the energy dependence of the classifiers. For that, we compared the *multilayer perceptron* with the method based on cuts and with a *multilayer perceptron* that does not includes the energy as a feature. We measured the signal efficiency for each method and for different energy ranges. Note, however, that training was made on the full range. Results are shown in Fig. 3b. While the efficiency of the cuts based method varies considerably with the energy, the performance of the classifier is more stable when the energy is used as feature.

6 Conclusions

We have presented a reconstruction algorithm for a preshower detector. We have shown that the algorithm is able to identify incident particles, separate overlapping particles and reconstruct the position and energy of the identified particles. Moreover, we have proposed a machine learning algorithm that can identify the incident particles when the reconstruction algorithm fails. We have found that the method based on a parametrized neural network outperforms the method based on cuts that is commonly used in *high energy physics*. By including the total energy of the incident particles we have reduced the dependence of the algorithm on the energy of the detected particles.

The reconstruction algorithm is general and can be extended to other calorimeter detectors. Moreover, the use of the reconstruction algorithm is not limited to *high energy physics* since these kind of detectors have important applications in other fields such as *nuclear medicine*. In future work we plan to try the algorithm on real data produced by accelerators.

Acknowledgements. This work was partially supported by the Research Project Basal FB 0821 and by Fondecyt Project 1170123.

References

1. Adam-Bourdarios, C., Cowan, G., Germain, C., Guyon, I., Kégl, B., Rousseau, D.: The Higgs boson machine learning challenge. In: NIPS 2014 Workshop on High-Energy Physics and Machine Learning, pp. 19–55 (2015)
2. Sadowski, P.J., Whiteson, D., Baldi, P.: Searching for Higgs boson decay modes with deep learning. In: Advances in Neural Information Processing Systems, pp. 2393–2401 (2014)

3. Racah, E., Ko, S., Sadowski, P., Bhimji, W., Tull, C., Oh, S.Y., Baldi, P., et al.: Revealing fundamental physics from the Daya Bay neutrino experiment using deep neural networks. In: International Conference on Machine Learning and Applications (2016)
4. Kuleshov, S., Brooks, W.: A preshower detector for forward electromagnetic calorimeters (2012). https://wiki.bnl.gov/conferences/images/b/bd/RD2012-13-Brooks_preshower_final.pdf
5. Lampl, W., Loch, P., Menke, S., Rajagopalan, S., Laplace, S., Unal, G., Ma, H., Snyder, S., Lelas, D., Rousseau, D.: Calorimeter clustering algorithms. Technical report (2008)
6. Sandhir, R.P., Muhuri, S., Nayak, T.K.: Dynamic fuzzy c-means (dFCM) clustering and its application to calorimetric data reconstruction in high-energy physics. Nucl. Instrum. Methods Phys. Res., Sect. A: Accel. Spectrom. Detect. Assoc. Equip. **681**, 34–43 (2012)
7. Morháč, M., Kliman, J., Matoušek, V., Veselský, M., Turzo, I.: Identification of peaks in multidimensional coincidence γ -ray spectra. Nucl. Instrum. Methods Phys. Res. Sect. A: Accel. Spectrom. Detect. Assoc. Equip. **443**(1), 108–125 (2000)
8. Berger, F., Bock, D., Clewing, G., Dragon, L., Glasow, R., Hölker, G., Kampert, K., Peitzmann, T., Purschke, M., Roters, B., et al.: Particle identification in modular electromagnetic calorimeters. Nucl. Instrum. Methods Phys. Res., Sect. A: Accel. Spectrom. Detect. Assoc. Equip. **321**(1), 152–164 (1992)
9. Awes, T., Obenshain, F., Plasil, F., Saini, S., Sorensen, S., Young, G.: A simple method of shower localization and identification in laterally segmented calorimeters. Nucl. Instrum. Methods Phys. Res., Sect. A: Accel. Spectrom. Detect. Assoc. Equip. **311**(1), 130–138 (1992)
10. Palma, F., Collaboration, C., et al.: Simulation studies of the expected proton rejection capabilities of CALET. In: 34th International Cosmic Ray Conference (ICRC2015), vol. 34 (2015)
11. Bellotti, R., Boezio, M., Volpe, F.: A particle classification system for the pamela calorimeter. *Astropart. Phys.* **22**(5), 431–438 (2005)
12. Cranmer, K., Pavez, J., Louppe, G.: Approximating likelihood ratios with calibrated discriminative classifiers. arXiv preprint [arXiv:1506.02169](https://arxiv.org/abs/1506.02169) (2015)
13. Baldi, P., Cranmer, K., Faucett, T., Sadowski, P., Whiteson, D.: Parameterized machine learning for high-energy physics. arXiv preprint [arXiv:1601.07913](https://arxiv.org/abs/1601.07913) (2016)
14. Agostinelli, S., Allison, J., Amako, K.A., Apostolakis, J., Araujo, H., Arce, P., Asai, M., Axen, D., Banerjee, S., Barrand, G., et al.: GEANT4 - a simulation toolkit. Nucl. Instrum. Methods Phys. Res., Sect. A: Accel. Spectrom. Detect. Assoc. Equip. **506**(3), 250–303 (2003)
15. Hoecker, A., Speckmayer, P., Stelzer, J., Therhaag, J., von Toerne, E., Voss, H., Backes, M., Carli, T., Cohen, O., Christov, A., et al.: TMVA-toolkit for multivariate data analysis. arXiv preprint [physics/0703039](https://arxiv.org/abs/physics/0703039) (2007)