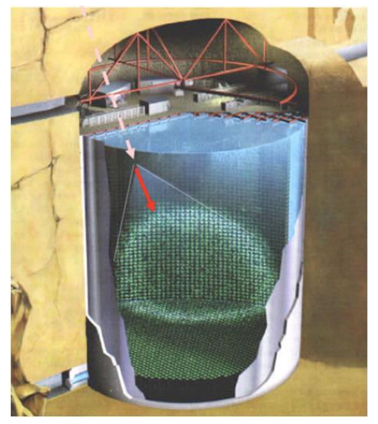
A Novel Algorithm to Reconstruct Events in a Water Cherenkov Detector



Mo Jia[1], Karan Kumar[1], Liam S. Mackey[2], Alexander Putra[3], Cristovao Vilela[4], Michael J. Wilking[1], Junjie Xia[5], **Chiaki Yanagisawa[1,3,*]**, and Karan Yang[6]

[1] Stony Brook University USA, [2] Rensselaer Polytechnic USA, [3] BMCC/CUNY USA, [4] CERN Switzerland, [5] University of Tokyo Japan, [6] Cornell Tech, USA

[*] Presenter for Poster #371

Introduction

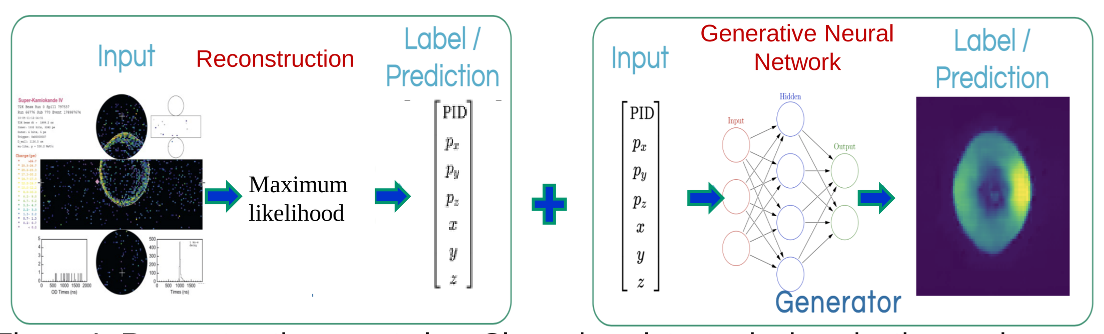


Figure 1. Reconstructing events in a Cherenkov detector is done by the maximum likelihood (ML) method (left), and modeling the ML function (right) is done by training a generative neural network in the new proposed method.

Traditionally in high energy physics, we try to reconstruct the event for a given event recorded by a detector and to obtain the properties of the event such as the particle ID, direction, energy, and event vertex (Fig.1). Note that in this poster I only talk about single-particle events detected by a Super-Kamiokande (SK)-like water Cherenkov detector. To reconstruct the properties of the particle detected, the current state-of-the-art non-neural-network based method for the SK event reconstruction is called fiTQun (Ref.1). It uses the maximum likelihood estimation algorithm.

FiTQun Likelihood Function

The log-likelihood function used by fiTQun is defined as:

$$-lnL(x) = -\sum_{j}^{unhit} lnP_j(unhit|x) - \sum_{i}^{hit}\left(ln[1-P_i(unhit|x)] + lnf_q(q_i|x) + lnf_t(t_i|x)\right) \quad (1)$$

where $P_j(unhit|x)$ is the probability that the j^{th} photomultiplier tube (PMT) does not register a hit, $f_q(q_i|x)$ is the probability density function (PDF) for the observed charge q_i in the i^{th} *hit* PMT, and $f_t(t_i|x)$ is the PDF for the observed time of the hit. x is the hypothesis on the origin of the particle that generated the Cherenkov ring pattern and it includes the particle ID, direction, energy and vertex position.

Event Reconstruction and Maximum likelihood Method

Event reconstruction is done by minimizing the -log-likelihood function $-lnL(x)$ over possible hypotheses. The hypothesis x that gives the minimum -log-likelihood is taken as the reconstructed event.

Drawbacks of FiTQun

To make the likelihood function tractable, it is factorized into several low dimensional components. For example, the contribution from Cherenkov photons that hit PMTs directly and that from photons reflected on the detector surface or scattered in the water are factorized. As the behaviors of the reflected and scattered photons depend on many factors such as details of the detector geometry, to have very useful simulated events which produce the required number of simulated events become unmanageably large.

Generative Neural Networks for Maximum Likelihood Reconstruction

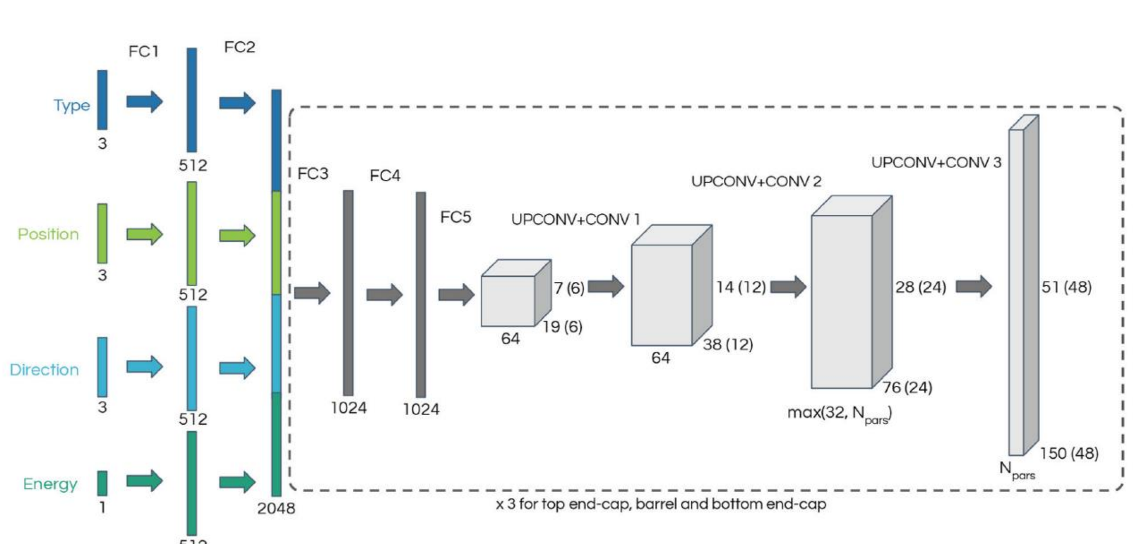


Figure 2. Architecture of the generative neural network used for this work

Network Architecture

Fig.2 shows the architecture of the generative neural network we use to model the likelihood function and we were inspired by Ref. 2. The inputs to the neural network (NN) are the particle ID (electron/muon), the vertex position, the direction and the energy. The outputs are charges and timings registered in all the PMTs. This neural network works as a fast event generator.

The network consists of five fully-connected (FC1-5) layers, three layers of a pair of a transposed (UPCONV) and a normal convolution (CONV) layer, and the last layer that produces charge and timing for each PMT with the desired dimensions for the barrel, top- and bottom-cap. Note that the depths/channels of the last layer depends on the parameters of functions to represent the charge and time distributions (see below). For more details, please refer to Ref. 3.

Loss Function

As the standard procedure, when a neural network is trained, the loss function is minimized to reduce the difference between the generated outputs and the training simulated events. We define the loss function in a similar fashion as the fiTQun's log-likelihood function Eq.(1) as follows:

$$Loss = -lnL = \sum_{i} -lnP_{unhit}(y_i) + \sum_{i_{hit}} -lnP_{qt}(q_{i_{hit}}, t_{i_{hit}}) \quad (2)$$

where the index i runs over all PMTs, and y_i is a label set to 1 if the PMT is not hit or 0 otherwise. The index i_{hit} runs only over the PMTs that register a hit in the event, and $P_{qt}(q_{i_{hit}}, t_{i_{hit}})$ is the PDF for observing charge $q_{i_{hit}}$ and time $t_{i_{hit}}$. $P_{unhit}(y_i)$ is the unhit probability for which we use the binary-cross-entropy loss from PyTorch with a Sigmoid function. The loss function is the sum of the contributions from the barrel and two caps.

Models for $P_{qt}(q,t)$

We model the PDF $P_{qt}(q,t)$ as a collection of Gaussian functions with and without the correlation between charge and time. We tried the number of Gaussian functions between 1 and 10 with and without correlation. In the case of no correlation,

$$-lnP_{qt}(q_{i_{hit}}, t_{i_{th}}) = -\Sigma_{i_{hit}}\left[\Sigma_j^N\left[ln(n_j) - ln\left(\sqrt{2\pi}\sigma_{q_j}\right) - \frac{(q_{i_{hit}}-\mu_{q_j})^2}{2\sigma_{q_j}^2} - ln\left(\sqrt{2\pi}\sigma_{t_j}\right) - \frac{(t_{i_{hit}}-\mu_{t_j})^2}{2\sigma_{t_j}^2}\right]\right] \quad (3)$$

where N is the number of Gaussians with σ and μ as the parameters for each Gaussian function, and n_j is the normalization parameter for j^{th} Gaussian function. The parameter values are optimized by training.

Training the Network and Its Performance

Training of the network was done with 75% of ~1×10^6 electron and muon simulated events, while leaving 25% for the testing of the network performance. The simulated datasets are uniformly distributed in kinetic energy (1 – 6,500 MeV for electrons, 150 – 6,500 MeV for muons), in vertex positions, and in directions. The training process is fed the entire datasets 50 times (50 epochs). In each epoch a batch of 200 randomly chosen events is used at a time until the full datasets are exhausted.

Fig.3 shows how the loss function values improve as a function of the epoch with 1, 3, 5, and 10 Gaussians for the charge q and time t PDFs in cases of charge-only (left), and uncorrelated (middle) and correlated q and t (right). Notice that the value of loss function improves with the number of Gaussians.

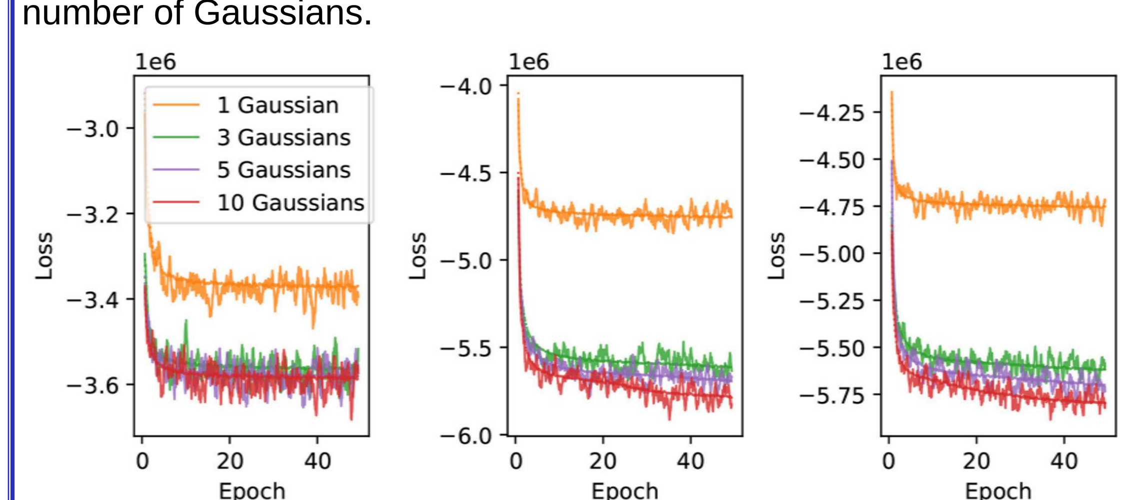


Figure 3. Improvement of the loss function with the number of epochs used: (right) q only, (middle) uncorrelated q and t, (right) correlated q and t.

The Cherenkov ring images of simulated Cherenkov electron and muon events in terms of hit probability are compared with respective predictions by the network are shown in Fig. 4. The predictions reproduce the simulated events well.

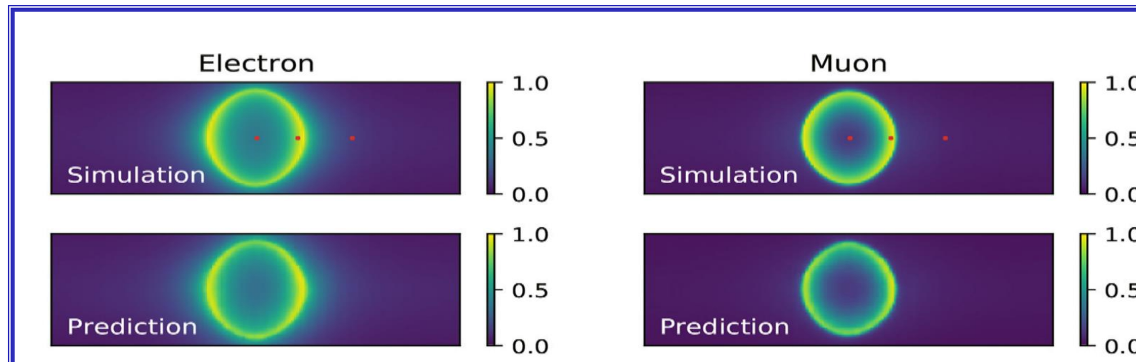


Figure 4. The hit probabilities of the simulated electron and muon events compared with the network predictions

Fig. 5 is an example of charge distribution with fitted multi-Gaussian function (1-10 Gaussians) for electron and muon of a PMT at the center of the ring marked with a red-dot in Fig.4.

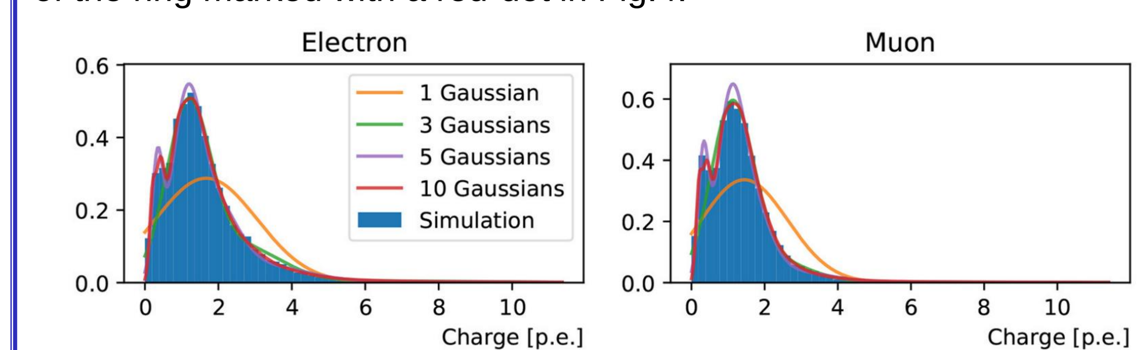


Figure 5. The charge distribution of electron and muon events compared with the predictions by single- or multiple-Gaussian functions

Results

Likelihood Scan as a Function of Particle Energy

For this presentation, we reconstruct the particle energy for a given event by scanning the log-likelihood function with a set of energy points and finding the minimum, while fixing other inputs such as the vertex position, direction, and particle ID. To see how well the network reconstruct the particle energy, we define the energy residual as: $\Delta_E = (E_{rec} - E_{true})/E_{true}$ where E_{rec} and E_{true} are the reconstructed energy and the true energy, respectively. Fig.6 shows the energy residual distributions of muon and electron events using the charge-only loss function with 1,3,5 and 10 Gaussians.

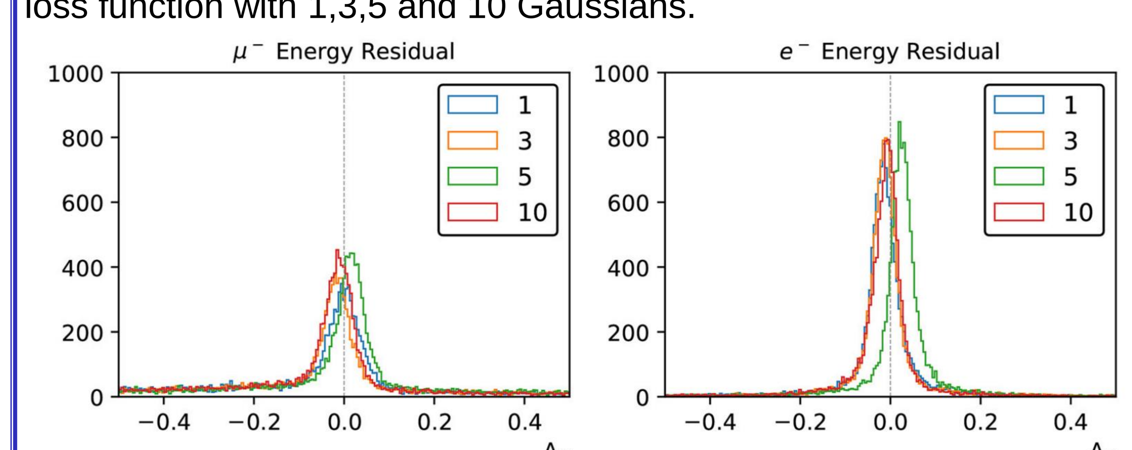


Figure 6. Δ_E distributions with 1,3,5 and 10-Gaussian PDF for charge only distribution

When a particles is produced near the boundary of the detector, a part of energy carried by the particle escapes detection especially in the case of muon. To assess this effect, we introduce a parameter called Towall that is the distance to the nearest detector boundary in the direction of the particle. We studied how the mean and the standard deviation of Δ_E change with three ranges of Towall with up to 10-Gaussian PDF of the PMT charge distribution. With the three ranges of Towall (0.0 – 200 cm, 200 – 500 cm, and 500 – 1700 cm), variations of the mean and standard deviation are shown in Fig.7 using 1-10 multi-Gaussian models for the charge PDF in the charge-only case.

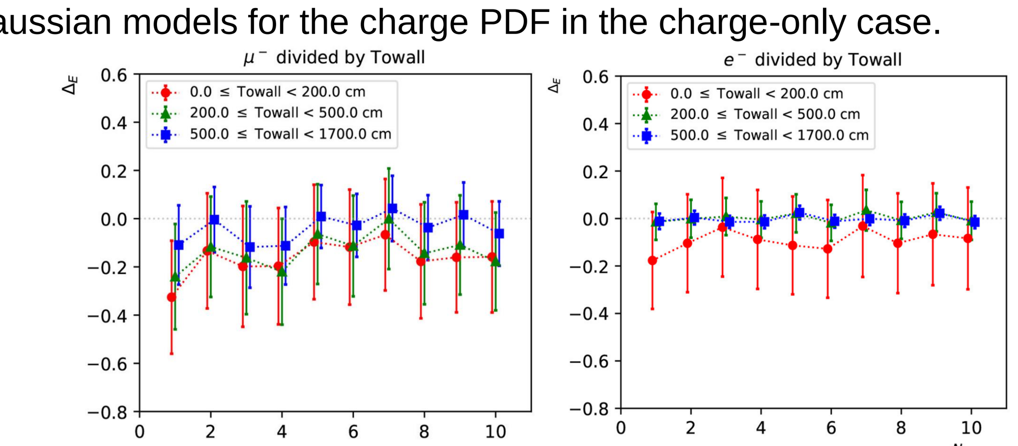


Figure 7. The means and standard deviations of the Δ_E distributions with three cuts on Towall using 1-10 multi-Gaussian models for the charge PDF

Particle Identification (ID): muon vs. electron

For particle ID between muons and electron, we compare two hypotheses for a given event whether it is a muon or an electron event. For an event, we run a muon and an electron neural network models that produce the respective loss function $-\ln L_\mu(x)$ and $-\ln L_e(x)$ where x represents the charge q, time t, reconstructed energy E_{rec} and the neural network parameters. The difference in the loss function is calculated $\Delta Loss = \ln L_\mu - \ln L_e$. If the event is electron (muon)-like, $\Delta Loss$ is negative (positive). Fig.8 shows the $\Delta Loss$ distributions with single- and 10-Gaussian models using the charge-only loss function.

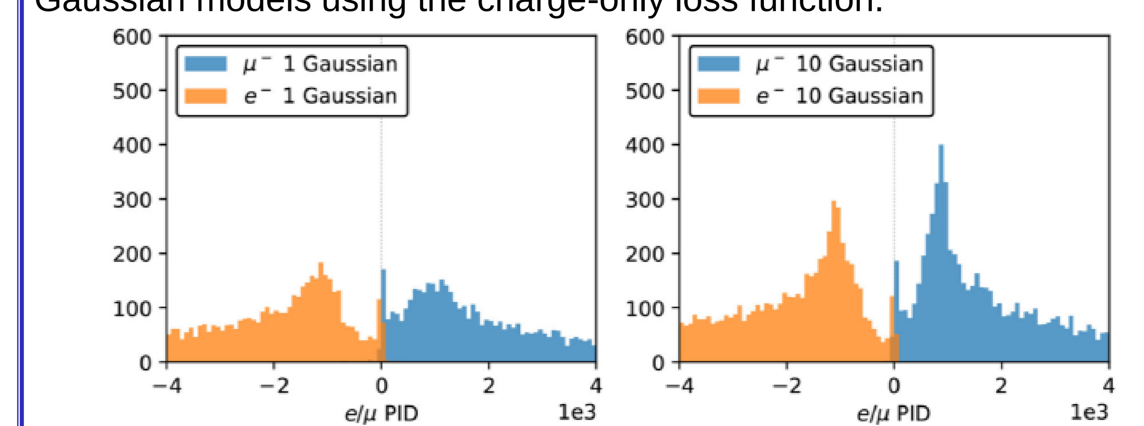


Figure 8. The distributions of $\Delta Loss$ with the single- and multi-Gaussian models using the charge-only loss function

As clearly seen, the separation between muons and electrons is excellent in general except for a spike near $\Delta Loss$ = 0. To see this is due to the particles close to the detector boundary, in Fig.9 we show the mis-identification probabilities of the true muons and electrons with the three cuts on Towall. Again using the events away from the detector boundary (Towall > 200 cm), the mis-identification probabilities decrease dramatically.

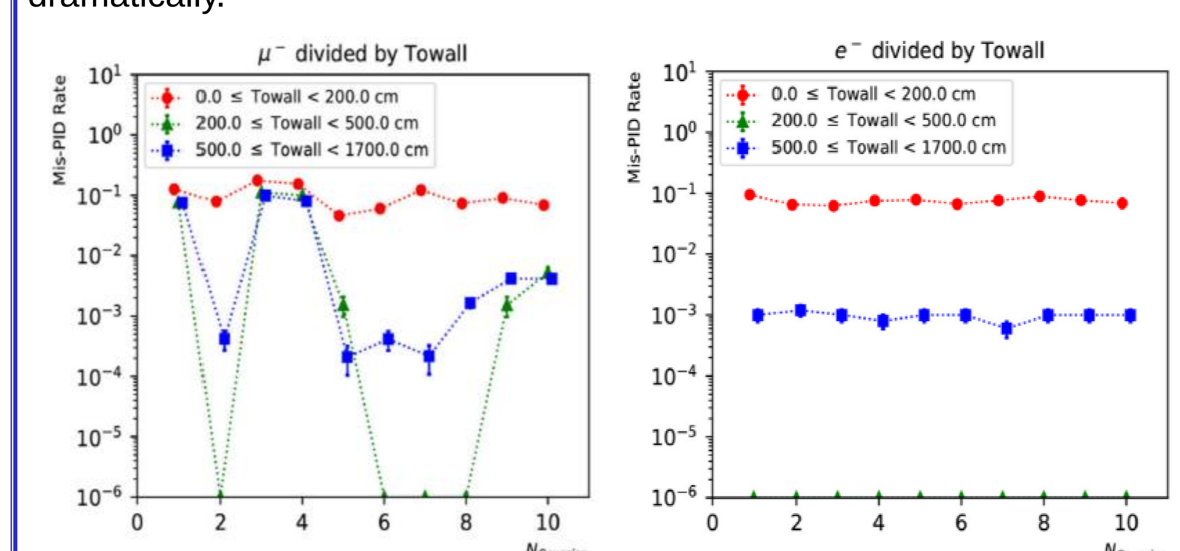


Figure 9. Particle mis-identification probabilities (rates) with 1-10-Gaussian PDFs for charge only distribution in the three Towall ranges (0 – 200 cm, 200 – 500 cm, 500 – 1700 cm).

Conclusions and Future Prospect

- I have shown the promising first results from a new way of reconstructing events detected by a Super-Kamiokande type water Cherenkov detector based on a generative neural network. **See Ref.3 for further details.**
- The next step is to try full reconstruction without fixing values for some inputs to the network. This is in progress.
- We are also looking into a possibility of distinguishing gammas from electrons.
- It is obviously important to try to reconstruct multi-particle events.
- There are other neural network algorithms such as a variety of Generative Adversarial Networks (GANs) that may improve the performance in reconstruction.

References

1. Jian, M et al. (2019). Atmospheric Neutrino Oscillation Analysis with Improved Event Reconstruction in Super-Kamiokande IV, PTEP 2019, 053F01
2. Dosovitskiy, A. et al. (2014). Learning to generate chairs with convolutional neural networks, arXiv.1411.5928, doi:10.1016/50168-9002(97)00048-X
3. Mo Jia et al. (2022). Maximum likelihood reconstruction of water Cherenkov events with deep generative neural networks, published in Frontiers in Big Data, 17 June 2022 at: https://www.frontiersin.org/articles/10.3389/fdata.2022.868333/full