Anomaly Detection on BESIII Electromagnetic Calorimeter using Machine Learning

Mingrun Li[1,2,*], *Xiaobin* Ji[1,2,**], and *Chunxiu* Liu[1,***]

[1]Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China
[2]University of Chinese Academy of Sciences, Beijing, China

Abstract. In the BESIII, unsupervised anomaly detection method based on autoencoders is applied to the CsI(Tl) electromagnetic calorimeter (EMC). This approach examines histograms of the electronic responses of all crystals, providing more accurate anomaly detection while requiring less personpower than traditional methods.

1 Introduction

The BESIII [1] at the BEPCII [2] electron-positron accelerator, which is located at Institute of High Energy Physics (IHEP), Beijing, China, is an experiment for the studies of hadron physics and τ-charm physics. It has collected large data samples in τ-charm region [3][4][5]. BESIII consists of a helium-based multilayer drift chamber (MDC), a time-of-flight system (TOF), a CsI(Tl) electromagnetic calorimeter (EMC) and a muon counter (MUC).

To ensure data quality and monitor the status of the detector, it is essential to conduct data quality monitoring and detector status monitoring. So far, there have been successful applications of machine learning for anomaly detection in detectors within the field of high-energy physics [6][7][8]. Applying machine learning methods for anomaly detection at BESIII presents a promising opportunity to enhance the efficiency of both data quality monitoring and detector status monitoring.

This paper presents an offline unsupervised anomaly detection method based on autoencoders applied to the BESIII EMC. The method analyzes histograms of electronic responses from all crystals, providing more accurate anomaly information while requiring less personpower compared to traditional methods.

2 Anomaly detection using autoencoder

An autoencoder is a type of neural network characterized by a special structure. It is widely utilized for data encoding, dimensionality reduction, and anomaly detection. The autoencoder's input dimension is equal to its output dimension, while the hidden layers are compressed. During training, the autoencoder aims to minimize a loss function between the input and output, and it will finally learn to encode the majority of data, while the minority of data will result in higher loss values.

*e-mail: mrli@ihep.ac.cn
**e-mail: jixb@ihep.ac.cn
***e-mail: liucx@ihep.ac.cn

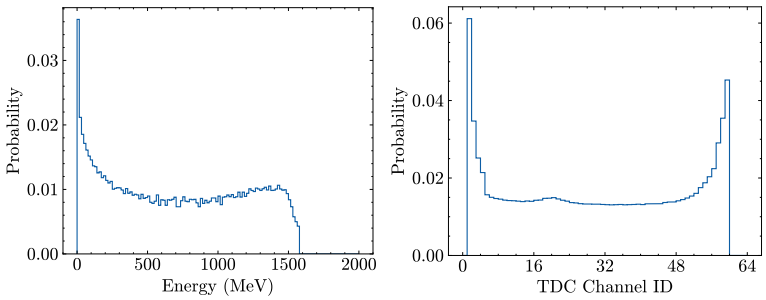

Figure 1. Left: Normalized energy distribution; Right: Normalized TDC channel distribution. Both histograms represent data from a normal crystal, populated using one day's worth of online Bhabha events.

With this characteristic, the procedure for anomaly detection using an autoencoder is as follows: 1) Ensure that anomalies constitute a minority in the data; 2) Train an autoencoder on the complete dataset; 3) Calculate the loss for each data point; 4) Identify data points with higher loss values as anomalies.

3 Data preparation

The electronics of the EMC crystals is designed with 1024 ADC channels and 64 TDC channels [1]. The ADC channels are allocated among three measurement ranges (low, medium, and high) [9], corresponding to different energy ranges. When an EMC hit is detected, the electronics selects the lowest measurement range that can accommodate the hit's energy to achieve optimal resolution.

To assess the status of each EMC crystal, histograms for hit energy and hit TDC channels are generated. Hit energy is calculated using ADC measurement ranges and channels, subsequently multiplied by a calibration constant. One day's worth of online Bhabha events is utilized to populate these histograms, resulting in 6240 energy and TDC histograms that reflect the status of 6240 crystals. These histograms are then normalized with bin number on the x-axis and used for anomaly detection. Fig. 1 illustrates normalized energy and TDC histograms for a normal crystal, filled using one day's worth of online Bhabha events.

3.1 Energy saturated crystal

When receiving high-energy hits, the electronics may become saturated. This is not an anomaly case, but crystals with saturated hits are grouped and processed separately. The saturated hits are disregarded, and the energy histograms are re-binned over the range [0, crystal-specific saturation cut-off], discarding the original binning scheme (Fig. 2). This approach ensures the autoencoder to focus solely on the shapes of the histograms, thereby excluding different energy cut-off points as determining factors.

To prevent statistical fluctuation being enlarged during the re-binning process, the bin number of re-binned histograms is chosen to half of the original energy histograms, as energy cut-off points in practice always exceed 1000 MeV – half of the bin range of the original energy histograms.

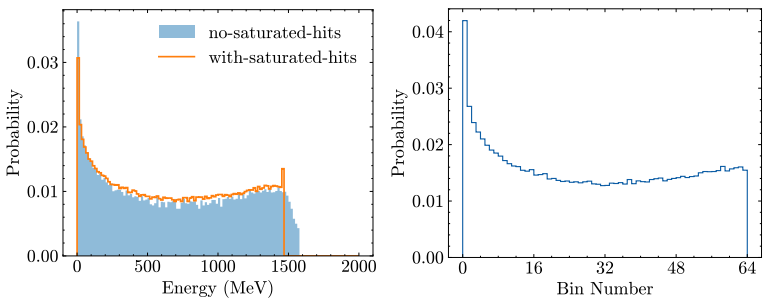

Figure 2. Left: Comparison of energy histograms for crystals with no saturated hits versus those with saturated hits. Saturated hits create a peak in the last bin. Right: The stretched energy histogram for the crystals with saturated hits in the left figure.

Table 1. Summary of histogram groups that will be processed independently

Group	N-Bins	Bin Range	Normalization
Energy, no-saturated-crystals	128	0-2000 MeV	Logarithmic
Energy, with-saturated-crystals	64	Crystal specific	Logarithmic
TDC, on-barrel	64	0-64 Channel	Linear
TDC, on-endcaps	64	0-64 Channel	Linear

3.2 Bhabha angular distribution

Due to the angular distribution of Bhabha events, crystals on the EMC endcaps receive significantly more hits than those on the EMC barrel. This phenomenon results in a higher peak in the TDC histograms for endcap crystals, as illustrated in Fig. 3. To mitigate this effect, the TDC histograms are divided into two groups: one for barrel crystals and one for endcap crystals. These two groups are processed independently using the same methodology. A summary of the histogram groups is provided in Table 1.

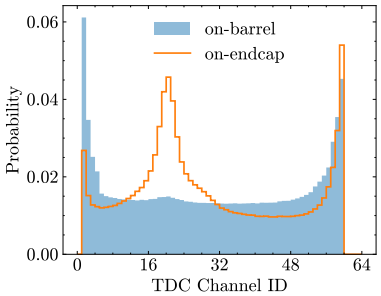

Figure 3. Comparison of normalized TDC histograms for barrel and endcap crystals.

4 Training

Anomaly detection on histograms is a relatively straightforward task, so a fully connected autoencoder is adopted. The Scaled Exponential Linear Unit (SELU) activation function is

Table 2. Model structure for different histogram groups

Group	Input	Hidden Layers	Output
Energy, no-saturated-crystals	128	[64, 32, 16, 8, L1, 16, 32, 64]	128 Softmax
Energy, with-saturated-crystals	64	[16, 8, L1, 16]	64 Softmax
TDC, on-barrel	64	[16, 8, L1, 16]	64 Softmax
TDC, on-endcaps	64	[16, 8, L1, 16]	64 Softmax

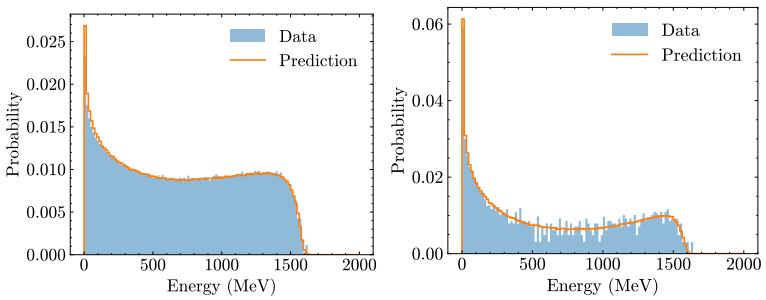

Figure 4. Energy histograms of an endcap crystal (left) and a barrel crystal (right), showing both the data and the autoencoder predictions.

utilized in the hidden layers, with L1 regularization applied following the most narrow hidden layer. The Softmax activation function is implemented in the output layer to assure the sum of the output values is 1. A summary of the structure of the autoencoder for different histogram groups is presented in Table 2.

Training is conducted using the Adam optimizer along with an early stopping mechanism and Mean Squared Error (MSE) loss function as loss function. The training process is halted when the loss ceases to decrease for 50 epochs, at which point the model with the lowest loss is saved.

5 Anomaly detection strategy

After training, the final losses for all crystals are calculated using the normalized histograms and the corresponding predictions from the autoencoders. The variations in the number of hits due to the Bhabha angular distribution introduce a bias in the loss calculations. Specifically, crystals with fewer hits tend to exhibit greater fluctuations in their energy histograms (Fig. 4), resulting in higher losses. To mitigate this bias, the losses are normalized by dividing by the median loss of the corresponding φ-direction ring.

For each histogram group, a Kernel Density Estimation (KDE) method is employed to estimate the probability density function (PDF) of the loss distribution. A probability density threshold is then established to cut off those outlier losses. The outlier cut-off point is defined as the first point where the PDF value downward intersects the probability density threshold. All crystals with losses higher than this cut-off point are tagged as anomaly candidates, as illustrated in Fig. 5.

Since there are no tags to indicate whether a crystal is normal or not, the probability density thresholds are empirically determined. The threshold selection strategy aims to maximize true outlier inclusion while keeping the quantity of false outliers low enough, so that manual review can be efficiently performed. Thresholds of all histogram groups are listed in Table 3.

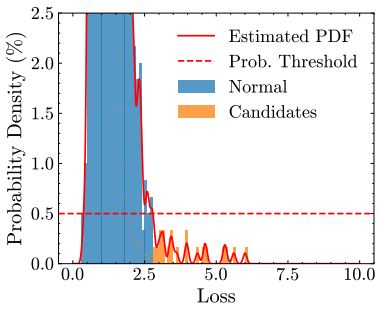

Figure 5. Estimated PDF alongside the probability density threshold and distribution of losses. The outlier cut-off point is defined as the right intersection point between the PDF curve and the threshold line.

Table 3. Probability thresholds for different histogram groups

Group	Threshold
Energy, no-saturated-crystals	0.005
Energy, with-saturated-crystals	0.005
TDC, on-barrel	0.001
TDC, on-endcaps	0.001

In practice, approximately 30 crystals are tagged as anomaly candidates based on their energy histograms, while around 20 crystals are tagged due to their TDC histograms. These anomaly candidates are then manually reviewed to ascertain whether they represent genuine anomalies.

6 Results

The anomalies detected from real data collected in one day, are depicted in Fig. 6 and Fig. 7. Given that one day's worth of data is sufficient for anomaly detection, the autoencoder can be trained daily, allowing for the detection of anomalies and tracking their status in time. Fig. 8 gives an example of tracking anomaly's status, repair and recovery.

Additionally, a comparison of the energy histograms of a normal crystal at different beam energies is presented in Fig. 9, demonstrating that the autoencoder can automatically adapt to variations in beam energy.

7 Conclusion

The autoencoder-based anomaly detection method presented in this paper has demonstrated significant advantages for detecting anomalies in the BESIII EMC. This approach is more automated, reduces personpower requirements, and increases accuracy. Several real anomalies have been detected and addressed, further proving the effectiveness of the method.

There are likely many more suitable applications for this technique, as its core principle involves analyzing the shape of histograms. For example, it could be effectively implemented in other detectors with similar data structures or applied to various types of data that can be represented as histograms.

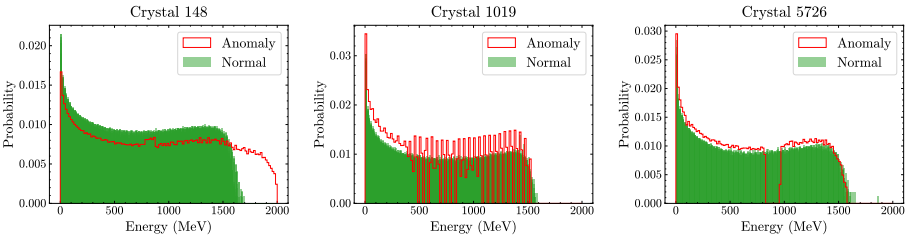

Figure 6. Energy histograms of anomalies in non-saturated-hit crystals. Red lines represent the histograms of the anomalies, while green bars are the histograms of neighboring normal crystals for each anomaly.

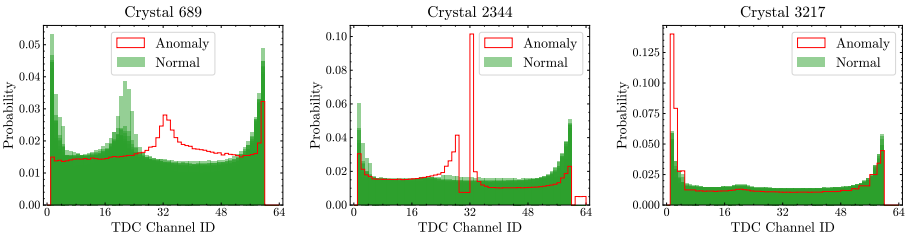

Figure 7. TDC histograms of anomalies in on-barrel crystals. Red lines represent the histograms of the anomalies, while green bars are the histograms of neighboring normal crystals for each anomaly.

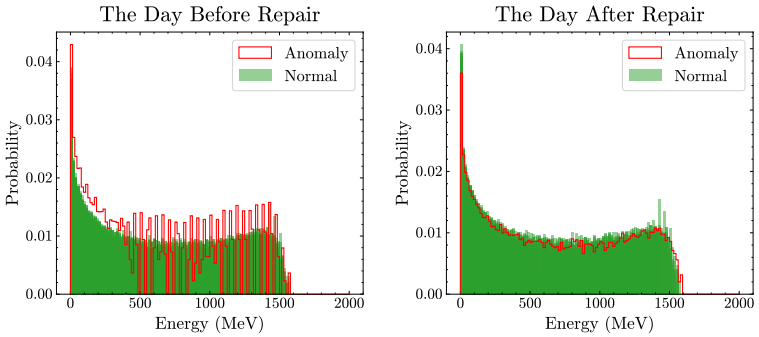

Figure 8. Crystal 1019 was identified as abnormal. It was then repaired by replacing the electronics. Red lines represent the histograms of crystal 1019, green bars are the histograms of its neighboring normal crystals.

8 Acknowledgments

This work was supported by the National Key R&D Program of China under Contracts Nos. 2020YFA0406300 and National Natural Science Foundation of China (NSFC) under Contracts Nos. 12475199, 11775245, 11875276.

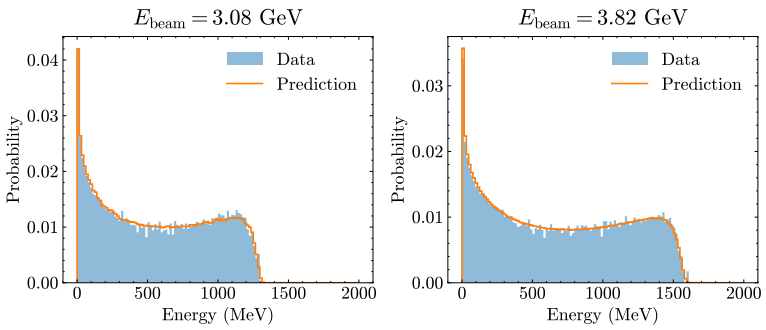

Figure 9. Energy histograms of a normal crystal at varying beam energies.

References

[1] M. Ablikim et al. (BESIII), Design and Construction of the BESIII Detector, Nucl. Instrum. Meth. A **614**, 345 (2010), 0911.4960. 10.1016/j.nima.2009.12.050

[2] C. Yu et al., BEPCII Performance and Beam Dynamics Studies on Luminosity, in *7th International Particle Accelerator Conference* (2016), p. TUYA01

[3] M. Ablikim et al. (BESIII), Future Physics Programme of BESIII, Chin. Phys. C **44**, 040001 (2020), 1912.05983. 10.1088/1674-1137/44/4/040001

[4] J. Lu, Y. Xiao, X. Ji, Online monitoring of the center-of-mass energy from real data at BESIII, Radiation Detection Technology and Methods **4**, 337 (2020). 10.1007/s41605-020-00188-8

[5] J.W. Zhang, L.H. Wu, S.S. Sun, J.H. Zou, C.X. Liu, Z.Y. Deng, Q.M. Ma, S.P. Wen, G.F. Cao, B.L. Zhang et al., Suppression of top-up injection backgrounds with offline event filter in the BESIII experiment, Radiation Detection Technology and Methods **6**, 289 (2022). 10.1007/s41605-022-00331-7

[6] A.A. Pol, G. Cerminara, C. Germain, M. Pierini, A. Seth, Detector monitoring with artificial neural networks at the CMS experiment at the CERN Large Hadron Collider, Comput. Softw. Big Sci. **3**, 3 (2019), 1808.00911. 10.1007/s41781-018-0020-1

[7] D. Abadjiev et al. (CMS ECAL), Autoencoder-Based Anomaly Detection System for Online Data Quality Monitoring of the CMS Electromagnetic Calorimeter, Comput. Softw. Big Sci. **8**, 11 (2024), 2309.10157. 10.1007/s41781-024-00118-z

[8] M.W. Asres et al. (CMS HCAL), Spatio-Temporal Anomaly Detection with Graph Networks for Data Quality Monitoring of the Hadron Calorimeter, Sensors **23**, 9679 (2023), 2311.04190. 10.3390/s23249679

[9] J. Chang, P. Wang, S. Gu, Z. Wang, J. Nie, The design of BESIII EMC charge measurement with three ranges, Nuclear Electronics and Detection Technology **29**, 718 (2009).