Anomaly Detection Based on Machine Learning for the CMS Electromagnetic Calorimeter Online Data Quality Monitoring

Abhirami Harilal[1,*], *Kyungmin* Park[1,**], and *Manfred* Paulini[1,***] (On behalf of the CMS Collaboration)

[1]Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

Abstract. Using a semi-supervised machine learning approach we present a real-time anomaly detection system based on an autoencoder used for online data quality monitoring of the CMS electromagnetic calorimeter operating at the CERN LHC. We introduce a novel method that maximizes the anomaly detection performance making use of the time-dependence of anomalies and the spatial variations in the detector response. The autoencoder-based system efficiently detects anomalies in real time and maintains a very low false discovery rate. We validate the performance of this novel system with anomalies from LHC collision data taken in 2018 and 2022. In addition, results are presented after deploying the autoencoder-based system in the CMS online Data Quality Monitoring workflow at the beginning of LHC Run 3 resulting in the system to detect issues that were missed by the existing system.

1 Introduction

The CMS detector has been taking proton-proton collision data produced by the CERN LHC for almost 20 years. The central feature of the detector is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid. A detailed description of the CMS detector, together with a definition of the coordinate system and relevant kinematic variables, can be found in Ref. [1].

The ECAL consists of 75,848 lead tungstate crystals, which provide a coverage of $|\eta| < 1.48$ in the barrel region (EB) and $1.48 < |\eta| < 3.00$ in the two endcap regions (EE+ and EE−) as displayed in Fig. 1. A total of 61,200 crystals provide a granularity that is 360-fold in ϕ and (2×85)-fold in η in the barrel. Each crystal has dimensions of 0.0174×0.0174 in $\Delta\eta \times \Delta\phi$ space, while each endcap is divided into two halves, with each comprising 3,662 crystals. Preshower detectors are located in front of each endcap detector. They consist of 2 planes of silicon sensors that are interleaved by 3 radiation lengths of lead.

An essential tool for the operation of CMS is the data quality monitoring (DQM) system [2] allowing it to record high-quality collision data. The current DQM utilizes a preliminary analysis of a subset of collected CMS data and consists of a software system which produces sets of histograms. CMS defines quality flags on these histograms through cut-based thresholds that are continuously monitored in the CMS control room by a DQM shifter who reports about any apparent irregularities observed. While the conventional DQM system has been very dependable over the years, increasing LHC collision rates and changing running conditions, together with aging electronics, expose failure modes that are newer and harder to predict with the present DQM.

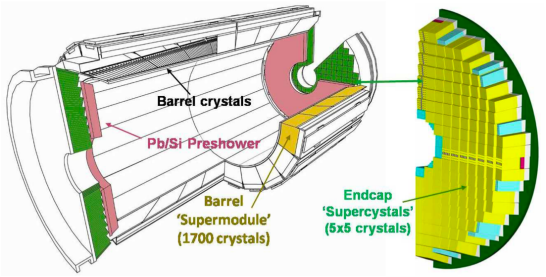

Figure 1. Schematic view of the ECAL showing the cylindrical barrel closed by the two endcap regions with one half endcap displayed.

The current ECAL DQM consists of two kinds of histograms: "Occupancy-style" histograms shown at the top of Fig. 2 filled with critical quantities from the real-time detector data and "Quality-style" histograms displayed at the bottom of Fig. 2. The latter are obtained by applying predefined requirements and thresholds to the occupancy-style histograms. The thresholds are typically derived from regular detector response. The quality histograms are displayed as easily identifiable maps using a color code scheme, for example, of red for "bad", and green for "good", or brown for "known problems".

*e-mail: aharilal@andrew.cmu.edu
**e-mail: kyungmip@andrew.cmu.edu
***e-mail: paulini@andrew.cmu.edu

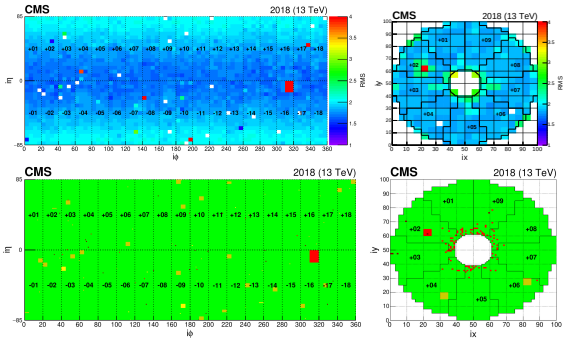

Figure 2. Example ECAL DQM diagrams indicating the distribution of RMS of the pedestal values in the barrel and EE+ (top). After a set of cuts is applied on the noise values of the top plots, the histograms at the bottom display the corresponding quality map for the two regions, drawn at a channel-level granularity.

2 Anomaly Detection Strategy Based on Machine Learning

Nowadays modern machine learning (ML) approaches are widely used in high-energy physics [3]. They provide new approaches for anomaly detection in particle physics searches [4]. We present a semi-supervised method utilizing anomaly detection based on an autoencoder (AE) [5] to supplement the ECAL DQM system. We train the network exclusively on certified good physics data, allowing it to learn the patterns of good data. This way anything that differs from the nominally learned patterns is detected as anomalous without the need to explicitly present anomalous data during training. We use data from CMS runs collected during LHC Run 2 in 2018 for training and validation of the AE network. The input to the AE are images of the occupancy map from a single time interval called "lumi-section" (LS) corresponding to an approximate time duration of 23 seconds of proton-proton collisions.

We utilize an AE network based on a computer vision technique and built the ML system on a convolutional neural network (CNN) architecture [6] that exploits ECAL data processed as 2D images. The encoder part of the AE compresses the input data into a lower dimensional representation, called the latent space. It contains a meaningful internal representation of the input data. The decoder follows and decompresses the encoded data back to the original image of the same dimensions, called reconstructing the image. A reconstruction loss ($\mathcal{L}$) measures how well the output matches the input. It is defined as the Mean Squared Error between the input (x) and the AE-reconstructed output (x') through $\mathcal{L}(x, x') = ||(x - x')||^2$.

Minimizing this loss function allows a network trained on good input to learn and reconstruct the images well. However, when presented with anomalous data, the AE returns higher loss in the anomalous region, forming the basis of the anomaly detection strategy. Using endcap images we illustrate the functioning of the AE in Fig. 3. The input occupancy image (top left) is given to the AE and a reconstructed image is output (top right). Next we calculate the Mean Squared Error on each tower and display it in the same coordinates as a 2D loss map. The anomalous region is highlighted with the higher loss as compared to the other parts of the image (see bottom-right of panel). We then calculate a threshold based on the anomalous loss values to flag the anomaly after applying certain post-processing steps (see Sec. 2.1). From the post-processed loss map with the applied threshold we create a quality plot (bottom left). It clearly shows towers with their loss above the threshold identified as anomalous (shown in red) and towers with loss below threshold are tagged as good shown in green, thus identifying the anomaly.

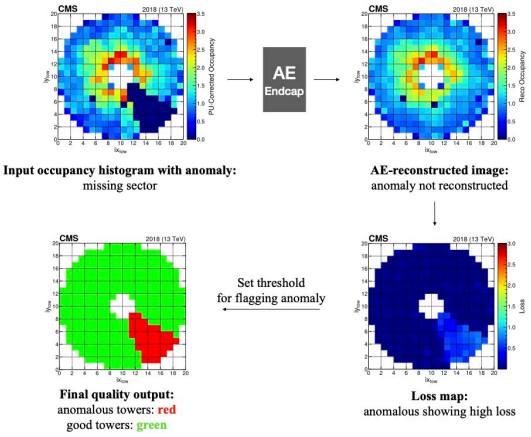

Figure 3. Diagram of strategy for AE-based anomaly detection.

2.1 ECAL Spatial Response and Time Corrections

To further maximize the anomaly detection efficiency and minimize the probability for false positive detections, we exploit the time-dependent nature of anomalies as well as spatial variations in the ECAL response to implement corresponding corrections.

Given the constant multiplicity of particle production at fixed rapidity, the number of particles per geometric interval increases for higher $|\eta|$ at a hadron collider. As a consequence we observe higher crystals occupancy at high $|\eta|$ compared to regions of low $|\eta|$ in both the ECAL barrel and endcaps. This difference in detector response is also visible in the AE loss map. Fig. 4 illustrates this effect with a missing supermodule consisting of 68 sets of 5×5 crystals. The top left plot gives the occupancy map with one supermodule showing zero occupancy. The figure to the right reflects the corresponding AE-reconstructed output, where the AE fails to reconstruct the anomaly. The tower-level loss map is calculated between output and input (see bottom left plot of Fig. 4) and shows high loss in the region of the anomalous supermodule. Due to higher average occupancy, towers at higher $|\eta|$ tend to have a higher loss than those at lower $|\eta|$. We normalize the loss by the average occupancy to mitigate this effect and obtain uniform loss in the anomalous region as indicated in the top-left plot of Fig. 4. Finally, we observe flat loss in the anomalous region after this "spatial response correction" shown in Fig. 4 (bottom right).

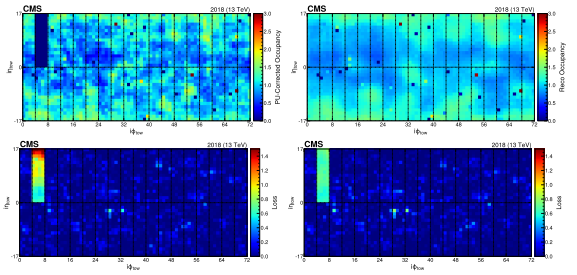

Figure 4. Top-left: Occupancy map with missing barrel supermodule. Top-right: AE-reconstructed occupancy map. Bottom-left: Loss map indicating the missing supermodule with higher loss at high $|\eta|$ due to differences in the detector response. Bottom-right: Loss map after spatial correction.

An additional correction exploiting the time-dependent nature of real anomalies brings a significant improvement in the AE performance . It is named "time correction" and based on the fact that real anomalies persist in time during consecutive LSs, while random fluctuations average out. Spatially corrected loss maps from 3 consecutive LSs are multiplied together at the tower level. The such time-multiplied loss map typically shows that the persistent anomaly of a real anomaly such as a missing supermodule is enhanced and random fluctuations from each LS are suppressed reducing false positives. We notice that a multiplication method turned out to be a better strategy for suppressing the resulting loss values rather than calculating an average.

3 Results

3.1 Anomaly Tagging Threshold and Performance Metric

The goal of the ML-based DQM system is to maximize the anomaly detection efficiency while minimizing the number of false positives. We tag an anomaly through a threshold obtained from validation data containing fake anomalies. We choose the threshold on the final post-processed loss map such that the loss values for 99% of anomalous towers are above the threshold. Showing the loss distribution from a zero occupancy tower scenario, this effect is illustrated in Fig. 5.

To assess the performance of the AE network, we use as a metric the False Discovery Rate (FDR) defined as the ratio of the number of good towers above the anomaly threshold over the total of good plus bad towers above threshold. Using the threshold chosen to catch 99% of the anomalies present in a dataset, the FDR value represents the fraction of false positive in all anomalies detected. This means, the FDR gives the ratio of good towers tagged as anomalous over all towers labeled as anomalous by the AE. A lower FDR means fewer false alarms during data taking and thus indicates better performance.

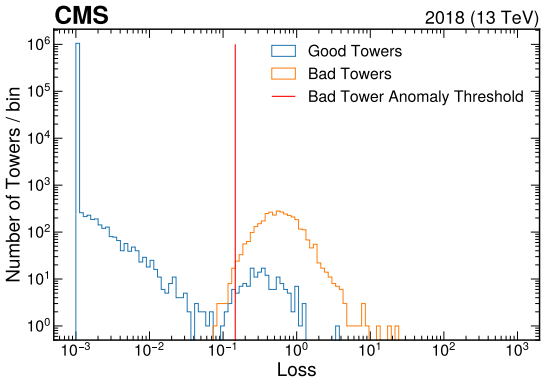

Figure 5. After spatial and time correction display of loss distribution for scenario of zero occupancy towers in EE−. A 1% anomaly threshold of all zero occupancy tower loss values is set.

Table 1. Summary of FDR for ECAL barrel fake anomaly scenarios using 99% anomaly detection threshold.

	FDR for 99% anomaly detection		
	Missing Supermodule	Zero Occup. Tower	Hot Tower
AE with no correction	0.036	0.51	0.028
AE after spatial correction	0.031	0.49	0.029
AE after spatial and time corrections	0.13%	0.041	< 0.01%

Table 2. Same summary of FDR as Table 1 but for endcaps.

	FDR for 99% anomaly detection					
	Missing Sector		Zero Occup. Tower		Hot Tower	
	EE+	EE−	EE+	EE−	EE+	EE−
AE with no correction	0.29	0.28	0.86	0.86	< 0.01%	< 0.01%
AE after spatial correction	0.018	0.022	0.11	0.14	0.02%	0.04%
AE after spatial and time corrections	0.06%	0.18%	0.014	0.044	< 0.01%	< 0.01%

3.2 Tests with Fake Anomalies

Adding artificial fake anomalies over good images, we study the performance of the AE-based DQM method with 3 distinct anomaly scenarios called single hot tower, single zero occupancy tower, and missing supermodule/sector. The corresponding FDR values are summarized in Tables 1 and 2, where we use anomaly tagging thresholds set for 99% anomaly detection optimized for each of the 3 scenarios. We observe the FDRs for the single zero occupancy tower scenario to always be higher than those for the single hot tower case independent of barrel and endcaps. The reason is that hot towers can in general be easier spotted. Compared to neighboring towers of average occupancy hot towers stand out with much higher occupancy.

Tables 1 and 2 show the effect of each consecutive correction on the FDRs. AE spatial corrections result in reduced FDRs in the missing supermodule/sector scenario as

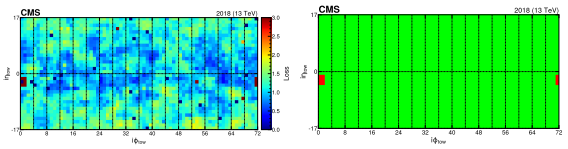

Figure 6. Input occupancy images from a 2018 run with hot towers and corresponding AE quality plots.

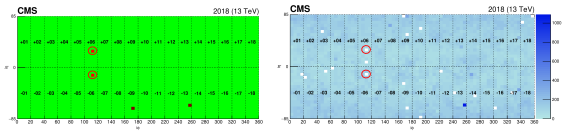

Figure 7. Left: New bad towers detected (circles) in 2022 run from ML quality plot obtained with E model of ECAL DQM. Right: Occupancy plot for one LS.

well as for single zero occupancy towers with occupancy values set to zero for the barrel/endcaps.

For towers with zero occupancy anomalies we find the loss values proportional to the towers' nominal occupancy without the correction indicating the loss is biased to larger values in higher $|\eta|$ regions (see, e.g., Fig. 4). Since the gradient in occupancy values across towers is more pronounced for the endcap regions, the spatial correction has a greater effect here than for the barrel. We note that the FDRs increase after the spatial correction for the case of hot tower anomalies since the hot tower loss is biased to be higher in the opposite direction, toward the lower $|\eta|$ region. However, these effects are mitigated by the time correction which greatly improves FDR values for all anomaly scenarios resulting in excellent final performance scores for both the barrel and endcaps.

3.3 Tests on Real Anomalies and Deployment

Next we study the AE performance on known anomalies taken from LHC data in 2018 and 2022. Fig. 6 shows the input occupancy images including anomalies with a region of hot towers and a zero occupancy tower from 2018 data on the left, while the AE quality output correctly identifies all the anomalous towers shown in red (see Fig. 6 on right). Interestingly this anomaly was not detected with the regular online DQM global quality plots in 2018, while the novel AE system detects this issue.

We deployed the AE-based anomaly detection system labeled MLDQM for the barrel in the CMS ECAL online DQM workflow in 2022 with the start of LHC Run 3 followed by the endcaps in 2023. New ML quality plots from the AE have been added to the ECAL DQM (see Fig. 7).

A trained PyTorch models exported to ONNX [7] and implemented in the CMS software framework using ONNX Runtime is used to accomplish model inference. So far, the MLDQM model has exhibited excellent performance with Run 3 data as illustrated in Fig. 7. Two red towers (circles) are indicated in supermodules EB+06 and EB−06. They correspond to zero occupancy towers as seen in the input occupancy map (Fig. 7 right). The AE quality plot also shows two known problematic towers.

4 Summary

Using a semi-supervised machine learning approach we have developed a production-level auto-encoder based anomaly detection and localization system for the CMS ECAL DQM indicating excellent performance. This work was just published in Ref. [8]. The ML based anomaly detection system described in this write-up can be further generalized and adapted to other subsystems of the CMS detector. In addition, other particle physics experiments can benefit from the presented system for their anomaly detection and data quality monitoring.

Acknowledgements

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid and other centers for delivering so effectively the computing infrastructure essential to our analyses. This research is supported in part by the U.S. Department of Energy, Office of Science, through DOE award DE-SC0010118.

References

[1] S. Chatrchyan et al. (CMS), The CMS experiment at the CERN LHC, JINST **3**, S08004 (2008). 10.1088/1748-0221/3/08/S08004

[2] V. Azzolini et al., The Data Quality Monitoring software for the CMS experiment at the LHC: past, present and future, EPJ Web Conf. **214**, 02003 (2019). 10.1051/epjconf/201921402003

[3] K. Albertsson et al., Machine Learning in High Energy Physics Community White Paper, J. Phys. Conf. Ser. **1085**, 022008 (2018), 1807.02876. 10.1088/1742-6596/1085/2/022008

[4] B. Nachman, Anomaly Detection for Physics Analysis and Less than Supervised Learning (2020), arXiv:2010.14554.

[5] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science **313**, 504 (2006). 10.1126/science.1127647

[6] Y. Lecun et al., Gradient-based learning applied to document recognition, Proceedings of the IEEE **86**, 2278 (1998). 10.1109/5.726791

[7] J. Bai, F. Lu, K. Zhang et al., Onnx: Open neural network exchange, https://github.com/onnx/onnx (2019)

[8] 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), arXiv:2309.10157. 10.1007/s41781-024-00118-z