Machine Protection & Electrical Integrity
TE–MPE

March 2025
Technical note 2025-01
philipp.hermueller@gmail.com
Christoph.Wiesner@cern.ch
EDMS Nr: 3252916

TECHNICAL NOTE

LHC Beam-Loss Detection and Classification Using Deep Learning

Philipp Hermüller[1], Christoph Wiesner[2], Vladimir Golkov[1,3], Fredrik Hülphers[1], Daniel Wollmann[2]
[1]TU Munich, Germany; [2]CERN, CH-1211 Geneva, Switzerland; [3]Munich Center for Machine Learning, Germany

Keywords: Beam Loss Monitors, Deep Learning, Machine Learning, Anomaly Detection, Classification

When high-energy particles at the Large Hadron Collider hit the beam pipe wall, machine equipment, or dedicated absorbers, they cause particle showers that get detected by beam loss monitors (BLMs). Machine protection experts use these measurements to label anomalies that occur around the moment of beam extraction. However, this requires expert knowledge and time. Therefore, we propose to automate this process with neural networks trained using supervised learning or unsupervised anomaly detection. The proposed neural network architecture is equivariant under permutations of BLMs but also uses one-hot-encoded BLM indices as inputs. This allows easily sharing feature extractors across different BLMs but also treating different BLMs differently when appropriate. The supervised-learning variant of our method classifies two well-known anomaly types – asynchronous beam dump tests and Unidentified Falling Objects – with an accuracy above 98% in a balanced test dataset. In contrast, existing methods that detect anomalies cannot distinguish different anomaly types. For unsupervised anomaly detection, we propose a multi-branch autoencoder whose latent features indicate the presence and “severity” of anomalies. Our methods aid the machine protection experts and the operations team in the control room, who need to analyse thousands of data sources, by indicting the likely origin of the beam abort.

Contents

1 Background **3**

2 Introduction **5**

3 Dataset **6**
3.1 Transform Data 7
3.2 Filter Data 8

4 Supervised Anomaly Detection **8**
4.1 General Architecture 8
4.2 Model: ASD-Pattern Detection 10
4.3 Model: UFO Losses Locally (Voting) 11

5 Unsupervised Anomaly Detection **12**

6 Results **14**
6.1 ASD-Pattern Detection 14
6.2 UFO-Pattern Detection 15
6.3 Abnormal Losses in Octant-7 15

7 Ablation Study **19**
7.1 Pooling Operation Choice 19
7.2 Data Normalization Choice 20
7.3 Missing Beam-Beam Kick 21
7.4 10 Hz Events in Beam 1 21

8 Conclusions **23**

9 Outlook **24**

A Octant Encodings of the Autoencoder **26**

1 Background

The nearly circular Large Hadron Collider (LHC) is a particle accelerator located at CERN in Geneva, Switzerland [Brü+04]. It is divided into eight octants, each consisting of two half-sectors. An octant is made up of a straight section and an arc at both ends. The LHC topology with its octants is shown in Figure 1. During LHC operation, two beams circulate in opposite direction within the two main beam pipes that only intersect at certain interaction points to allow for particle collisions. When a beam abort is triggered, both beams are extracted from the main beam pipes into two separate beam dump lines using fast kicker magnets. There are two main reasons why the beams are dumped. First, a person in the control room manually initiates a beam abort; reasons are amongst others to refill the LHC with new beam to increase the number of particle collisions. Second, if an anomaly occurs within an LHC system, the beam is dumped to protect the machine equipment and/or the the detectors.

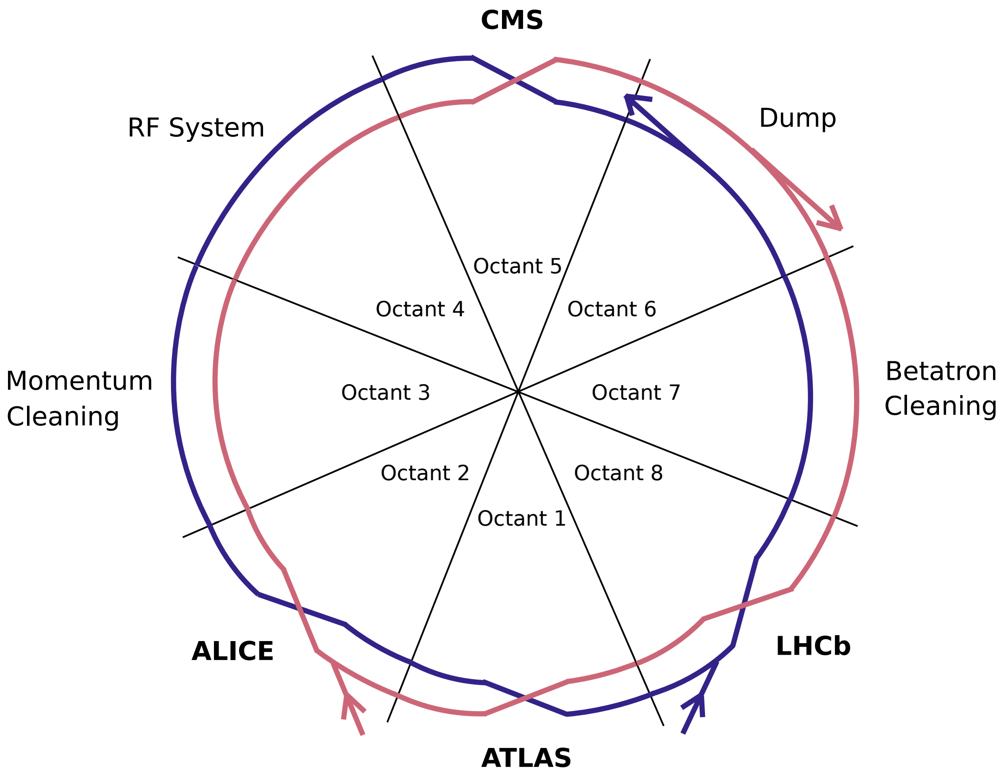


Figure 1: The topological layout of the LHC. It is divided into octants and each octant has a separate functionality. Octant-3 and 7 clean the beam, and Octant-4 accelerates the beam. In Octant-6 the beam extracted into dedicated extraction lines. The other octants contain detectors for particle collision events.

Beam Loss Monitors. Beam losses occur when high-energy particles are drifting from the beam path and interact with dedicated absorbers, machine equipment, or the beam pipes. In such cases, particle showers and ionizing radiation is created. They are observed using beam loss monitors (BLMs) that are installed at selected places all around the LHC ring and along the beam dump lines. In our case, we focus on BLMs that use *Ionization Chambers* (ICs) and *Little Ionization Chambers* (LICs) as detector, which presently amount to 3956 BLMs. Within a BLM, the ionizing radiation causes a directly proportional current flow that is passed to a data acquisition card [Eff+07] [Vig+12] for further processing. This card uses a Current-to-Frequency converter and counts the peaks in a 40 µs integration time window before converting it to a digital value. This value is called Running Sum 01 (RS01) and has the shortest integration time window. These RS01 values in units of 'bit' are directly proportional to the beam losses and have a value domain of [0, 262 600]. From the formerly mentioned basic integration time window, 12 running sums are constructed with different integration intervals. The running sum with the shortest integration time is called RS01 and the one with the largest is called RS12. When a beam dump occurs, a time series of 25 600 RS01 values is recorded; forming a window of 1.024 s (= 25 600 · 40 µs) around the moment of the beam dump. The maximum of this RS01 time series is referred to as RS01-max. From previous studies [Hul22], it was demonstrated that the RS01-max holds enough information in connection with context variables for the task of detecting and classifying beam losses.

Context Variables. Context variables refer to machine configuration settings as well as measured properties. For this work, we consider the following beam-specific properties:

- *Beam intensity* i.e. the total number of charges or particles stored in beam 1 or 2 [Gas+16, p. 32],
- *Abort gap population* of beam 1 & 2. Because the beam extraction kickers need a ramp-up time to reach their target field, a predefined time window of 3 µs is intentionally left empty within the beam filling pattern to avoid particle losses during the kickers' rise time. This time window is called the abort gap and the number of particles (charges) within it is the abort gap population.

BLMs record different loss distributions depending on their location within the LHC ring. We consider BLMs that usually observe beam losses within the noise level during nominal dumps as especially important for this work. During nominal dumps, these BLMs observe beam losses below 120 bit, which is mostly noise, regardless of the beam intensity and abort gap population. This makes them valuable for locating specific types of anomalies like UFOs.

Anomalies. An anomalous beam dump is defined as a beam dump where one or multiple BLMs have observed beam losses that have a low conditional probability given the used context variables. These anomalies can be an unexpected distribution of beam losses between BLMs or a high sum of beam losses, which can occur just before or during the beam extraction. Any beam dump that is not anomalous according to the prior definition is normal and referred to as a *nominal dump*. In this work, we consider three well-known types of anomalies:

- *Asynchronous beam dump test* (ASD test) [Wie+18]. ASD tests are manually initiated tests that create beam losses mainly in octant 6 and octant 7 where the beam dump line and the collimation or beam cleaning system are located, respectively.
- *Unidentified Falling Object* (UFO) [Lec+22] [Bél+22] [Mir+19] [Lin+20]. UFOs are dust grains interacting with the beam, leading to particle scattering that results in unexpected beam losses. These losses mostly appear in BLMs that are in close proximity to one another.
- *10 Hz Events*. During these events, the beam starts oscillating with a frequency of about 10 Hz, potentially causing beam losses in the collimation system in octant 7.

2 Introduction

After each beam dump, machine protection experts analyse multiple data recordings such as beam loss monitors to validate the correct functionality of the machine protection system and check if an anomaly occurred before or during the beam dump. For example, if a UFO occurs, they label the beam dump accordingly and also document the UFO's location and beam loss behaviour. Overall, "manual beam-loss analysis is time-consuming, not easy to standardise, and requires expert knowledge" [Wie+23]. Therefore, this work aims to automate the process by utilizing Deep Learning. The goal is a tool that automatically detects anomalies, classifies known anomalies, and gives a standardized representation of the beam dump analysis. This proposed tool would enable the operations team in the control room by indicting the likely origin of the beam abort.

A wide variety of Deep Learning techniques exist for this task e.g. supervised classification, anomaly detection by reconstructing normality [Bau+21], unsupervised clustering with auto-encoders [Ayt+18], and supervised similarity learning [HA15] amongst others. Related work [Hul22] [Wie+23] was already performed in this field for classifying beam aborts into OK and NOT-OK. Their approach creates a threshold value for each BLM based on historical data and classifies a beam abort as NOT-OK if two or more BLMs are above threshold. This technique only classifies events into anomalous or non-anomalous and does not classify the occurred anomaly type. Therefore, we propose a neural network architecture for the classification of ASD and UFO events, which is equivariant with a variable input size. Additionally to being able to classify UFO dumps, our defined model is also able to locate the longitudinal position of the UFO. Furthermore, we propose an octant-feature-based autoencoder for the classification of severe beam losses within octant 7. However, not only does the neural network provide a standardised value for beam losses within said octant, but it also provides some limited insight into beam losses at other octants. In summary, our contributions are

- An equivariant neural network architecture with variable input size for ASD and UFO dump classification as well as UFO localization.
- An octant-feature-based autoencoder that classifies severe beam losses within octant 7 as well as providing some limited insight into other octants. The results look promising to motivate future work to classify even more anomaly types and provide a more in-depth analysis of beam dumps.
- A study on the limits and possible extensions of our proposed networks.

3 Dataset

The dataset consists of proton beam dumps at top energy from LHC Run-2 (year 2016-2018, excluding the commissioning year 2015) and the start of LHC Run-3 (year 2022) of the LHC. Each BLM dataset sample consists of 3956 values, which are the RS01-max values for each connected BLM. According to the 10-times rule [HRS11], the dataset would need 39 560 beam dumps for stable training; however, the dataset consists of only 1339 beam dumps — 1083 Run-2 + 256 Run-3. Furthermore, only 1083 beam dumps (Run-2) can be used for training while the Run-3 beam dumps are used as the test set. Additionally, the dataset has a heavy class imbalance favouring the *non-anomalous* beam dumps with only a few samples representing UFOs, which further complicates stable training. This heavy class imbalance is visualized in Figure 2. To differentiate between NOT-OK and OK beam dumps as well as the severity of anomaly, beam dumps from Run-3 also have an *expert assessment factor* that a machine protection expert manually labelled.

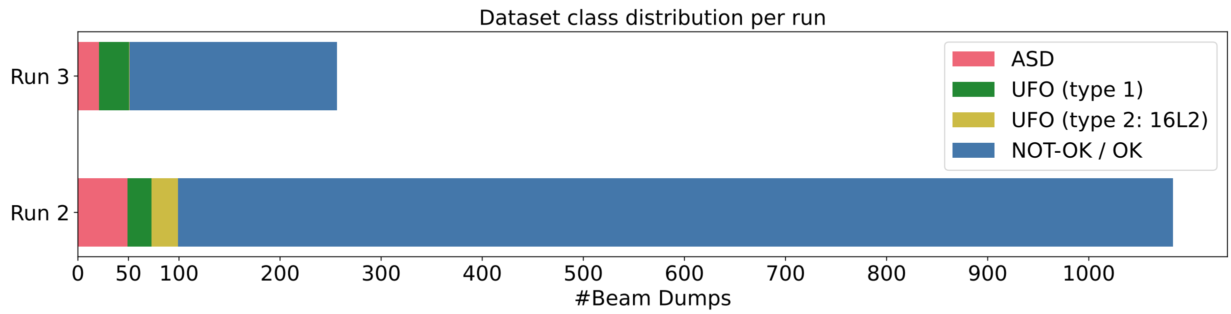


Figure 2: Stacked bar chart of the dataset's class distribution per run. The proton high-energy beam dump dataset has a heavy class imbalance with ASD test and UFO (type 1 & 2) events being under-represented in both runs. Furthermore, the training set (Run-2) only contains 1083 samples. In order to cope with this heavy class imbalance and the small dataset size, we make use of similarity learning and lightweight model architectures with few weights, as discussed in section 4.

3.1 Transform Data

In order to facilitate the learning process of the neural network, the RS01-max data values are transformed to combat the imbalance of their value distribution and use the value range more effectively. The transformation is given by

$$t(x) = \ln(\max(r) + c), \tag{1}$$

where the max of the RS01 time series $r \in \mathbb{N}^{25\,600}$ is log-scaled with a small constant $c = 1$ to avoid taking the log of zero. Most BLMs record RS01 values in the lower value range. Normalizing this skewed value range based on the RS01 saturation value of around 252 000 bit into the range $[0, 1]$ would result in losing the precision of lower values and having a mostly unused value range for higher values. However, log-scaling these values keeps the differences between RS01-max values while reducing the scale of this skew (see Figure 3).

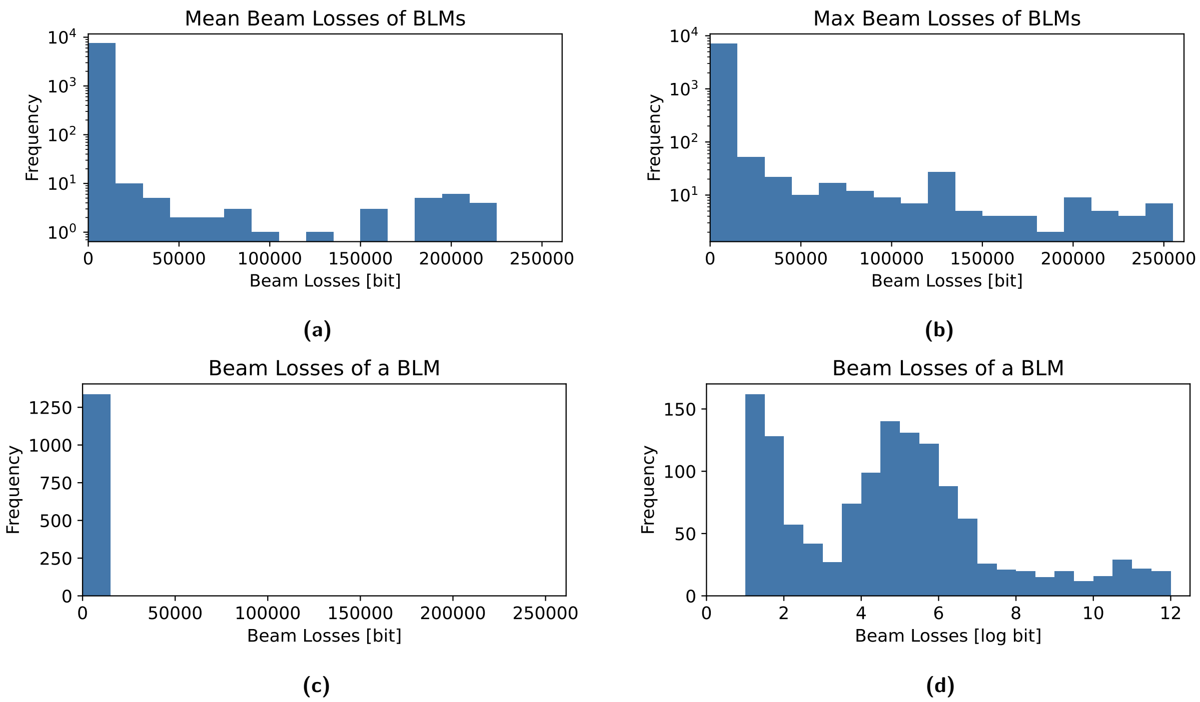


Figure 3: Data distribution histograms and influence of log-scaling the data. Most of our data is limited to the lower ranges which can be seen in (a) the histogram of BLM *mean* losses and (b) the histogram of BLM *max* losses. Furthermore, we show histograms of beam losses for the BLM 'BLMTI.06L7.B1E10_TCSG.A6L7.B1' with (c) losses non-scaled in bits and (d) losses in log-scaled bits. The non-scaled values are too skewed towards the lower range, leading to the entire value range not being used effectively. Applying the log on these values reduces the scale of this skew while keeping the differences between the losses, which improves the data range usage as input for the model.

3.2 Filter Data

Some beam dumps within the dataset are removed because they contain values that do not facilitate the learning process of the model but instead clutter the input or dataset with unnecessary data or impossible values. We filter the data according to the following criteria:

1. **Remove Duplicates.** Beam dumps need a time difference of at least 1h from the previous dump. This is the minimum time required between a previous beam dump and a new beam to reach top energy again.

2. **Remove Unphysical Values.** Filter out beam dumps with an *abort gap population* above 10^{14} charges.

3. **Filter to Specific BLMs Types.** Only use BLMs with an IC or LIC detector.

4 Supervised Anomaly Detection

Two separate neural networks are being used for ASD and UFO pattern detection. Both are based on the same architecture. The difference between these models is the used data for training and their class split as well as the employed global pooling operation.

4.1 General Architecture

Our proposed encoder architecture, as shown in Figure 4, is equivariant, handles variable input length, and consists of four parts: the BLM-index encoder, BLM-value encoder, BLM-scenario encoder, and a global pooling operation with a post-processing layer. The encoders do not have many layers and weights to avoid overfitting the small dataset. The layers of each encoder are listed in Table 1. Because our architecture can handle various input lengths, we do not need to use padding or flags for representing missing data within a beam dump. This alleviates our model in training with a tiny dataset size.

A beam dump from our dataset consists of tuples, one per BLM, consisting of the RS01-max and its unique BLM ID. The BLM-ID is a unique identifier with a random order and no relation to the order within the LHC ring. The BLM-Index Encoder turns the One-Hot Encoded BLM-ID into a 10-dimensional vector. In this way, it is learning the class and properties of specific BLMs. The BLM-Value Encoder increases the dimensionality of the RS01-max value in order for the next encoder to be able to rely more on this value; this is similar to a learned positional encoding [Dev+18], where the dimensionality is being increased. The BLM-Scenario Encoder encodes the concatenation of the previous two encoders into a four-dimensional scenario vector, resulting in a 4-dimensional vector for each BLM. Finally, a global pooling operation across the dimension of the BLMs is applied over all B BLM-scenario encodings reducing the dimensionality from $\mathbb{R}^{4\times B}$ to $\mathbb{R}^4$. This pooling operation is necessary to remain equivariant and transform the local BLM encodings to a global encoding. For fine-tuning purposes, the previous encoding will pass through a post-processing layer, resulting in the final similarity encoding.

Loss Function The *Triplet-Loss* [WS09] is being used for training the network because besides learning similarity clusters it also creates a ranking between beam dumps. Furthermore, and most importantly, per definition it uses a class-balanced batch, which is critical for our imbalanced dataset. In this case, the learned ranking would be from a clean dump to a beam dump with a specific anomaly. The Triplet Loss requires a positive and negative sample F_i^p and F_i^n in respect to the anchor sample F_i^a, where a positive sample has the same class as the anchor sample. The Triplet Loss is defined as follows:

$$L_T = \sum_{i=1}^{N} \max\left(\|F_i^a - F_i^p\|_2^2 - \|F_i^a - F_i^n\|_2^2 + \alpha, 0\right), \tag{2}$$

where α is some margin and N is the number of beam dumps.

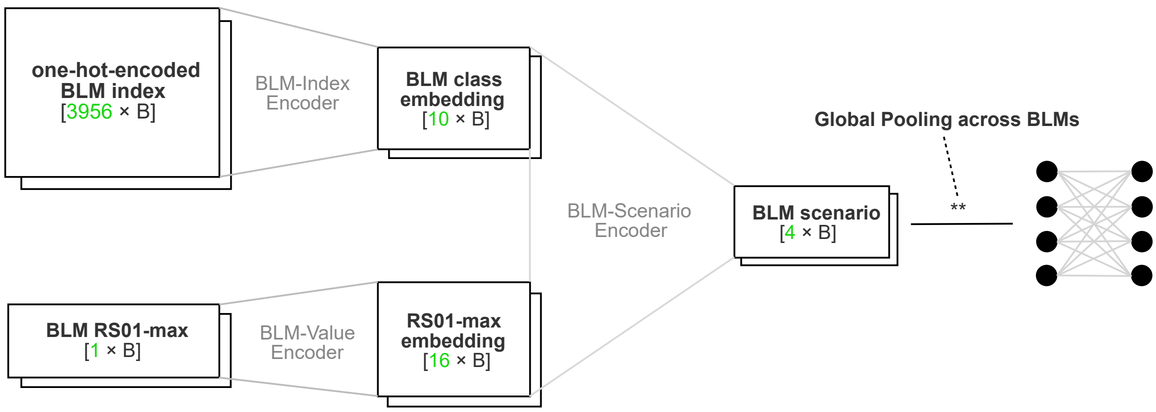


Figure 4: Our proposed encoder architecture for ASD and UFO classification. The encoder is equivariant under permutations of BLMs and provides interpretable intermediate results. Dimensions in green are the feature size and B represents the number of input BLMs. The output of the BLM-scenario encoder passes through a global pooling operation across the BLMs that is different depending on the aim of the training and can be either *mean* or *max*. The resulting four feature values that are post-processed by a Multi-Layer-Perceptron (MLP) represent the final similarity encoding between ASD and non-ASD or UFO and non-UFO.

Encoder	Channels-In	Channels-Out
BLM-Position Encoder	3956	1024*
	1024	256*
	256	10
BLM-Value Encoder	1	8
	8	16*
BLM-Scenario Encoder	26	4

Table 1: Layer composition of all three sub-encoders. Our network uses 1d-convolutions with a kernel size of one for all layers in all three encoders. Furthermore, it uses a ReLU activation after each layer as well as a Dropout of 10% for some layers, which are marked as (*) within the table.

4.2 Model: ASD-Pattern Detection

Asynchronous beam dump tests are tests that generate relatively high beam losses. Therefore, creating a model for detecting ASD tests serves as an important validation for the architecture to handle the input data format as well as the tiny dataset size.

Architecture. The used global pooling operation is the *mean* because during an ASD test many BLMs observe higher losses, which is different from other anomaly events that only affect a few BLMs. The post-processing layer also uses the beam intensity and abort gap population of each beam as context variables — see Figure 5.

Training. For a more stable training with fewer weight-update steps than usual we simulate a batch size of 4. As the input has a variable size, a batch of multiple samples of different shapes cannot be created. Therefore, the gradients are accumulated over four batches of batch-size one before backpropagation. For the Triplet-Loss, we consider the two class sets $\mathcal{S}$ and $\mathcal{R}$ defined as

$$\mathcal{S} = \mathcal{A}, \tag{3}$$

$$\mathcal{R} = \{b \in \mathcal{B} \,|\, b \notin \mathcal{A}\}, \tag{4}$$

where $\mathcal{A}$ and $\mathcal{B}$ are the sets of all ASD-test samples and all data samples respectively.

Data Normalization. The log-scaled RS01-max values, which are between zero and $\ln(252\,100) \approx 12.5$ are divided by 12.5 to be in the range $[0, 1]$.

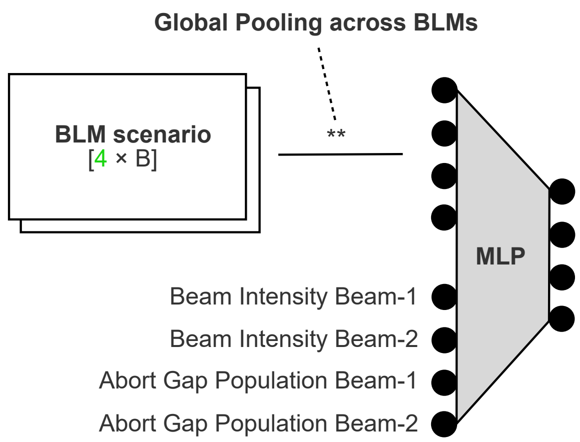


Figure 5: The Multi-Layer-Perceptron (MLP) used for post-processing in the ASD-Pattern Detection model. The ASD model also uses the *context variables* as additional input for the post-processing as shown here.

4.3 Model: UFO Losses Locally (Voting)

This model follows the general idea that UFOs show abnormally high losses in only two to five BLMs of the approximately 4000 available BLMs. With only very few input values showing abnormal behaviour, the *mean* global pooling operation is not the best fit for this scenario.

Architecture. The used global pooling operation is the *max* function because only a small fraction of the BLMs show UFO losses. By using the *max* function as a pooling operation over all BLMs, the network learns a voting system where each BLM 'votes' if abnormally high losses typical for UFOs are observed.

Training. Instead of simulating a higher batch size as done for the ASD Pattern Detection model, the UFO model trains on batches of size one as the model needs more weight-update steps. If we simulated a higher batch size, fewer weight-update steps are performed that leads to the model not learning long enough. Increasing the number of epochs, however, would lead to overfitting because of our tiny dataset. Still, each triplet sample already uses data from three beam dumps of mixed classes, which avoids a class-biased weight update. For the Triplet-Loss, we consider the following two class sets $\mathcal{S}$ and $\mathcal{R}$ defined as:

$$\mathcal{S} = \mathcal{U}, \tag{5}$$

$$\mathcal{R} = \{b \in \mathcal{B} \,|\, b \notin \mathcal{A}, b \notin \mathcal{U}\}, \tag{6}$$

where $\mathcal{A}$, $\mathcal{U}$, and $\mathcal{B}$ are the sets of ASD-test samples, UFO samples, and all available data samples respectively.

Data Normalization. The ln-scaled BLM values, which are between zero and $\ln(252\,100) \approx 12.5$ are divided by 6 to be in the range [0, 2.08] for the neural network training phase. Other data ranges have been tested as well, however, this range yielded the best results. During inference, the data normalization method is changed to normalize based on normal beam dumps (NOT-OK and OK) of the training data. This simulates other outlier detection approaches where inputs contain values out of distribution. The BLM-Position Encoder as well as the BLM-Value Encoder are not affected by this change in normalization because the former is independent of the value and the latter is a learned positional encoding that retains the linearity of the input. The important change is within the BLM-Scenario Encoder. By training the network with a specific normalization over the entire value range, the network is able to learn that specific BLMs always observe low losses while other BLMs observe losses of varying degrees. With the change in normalization at inference time, the BLM-Scenario Encoder can better differentiate between *Always Low-Loss BLMs*, which are important for UFO-Detection, and other kinds of BLMs.

Post-Processing. Post-processing is required because the proposed architecture does not support information on the LHC layout and BLM positioning. Therefore, the characteristic property of UFOs — high beam losses in neighbouring cells within the same octant — is unknown to the model. Using topological information as further input to the model would result in more information that the model would need to interpret. However, because of the small dataset size, it cannot handle too much input information. To filter out wrong predictions due to the missing topology information in the model, two criteria for the UFO votes are defined: there must be at least 2 votes above the threshold, and 66% of these votes need to be within the same octant. The BLM responsible for the highest UFO vote conforming to the octant-requirement is regarded as the original location of the UFO.

5 Unsupervised Anomaly Detection

We employ a variation of an autoencoder trained in an unsupervised setting to classify rare anomalies and obtain a standardized representation of each beam dump.

Architecture. The approaches from section 4 used a group operation over all BLM encodings which avoids information exchange between various BLMs. However, using an MLP with nearly 4000 input nodes and a small dataset would lead to overfitting. To facilitate the learning process, features of octants are encoded first. This means the network encodes all $N = 3956$ input features, which are the RS01-max values $y_{0 \leq i < N}$, by encoding features from each octant of the ring separately. Afterwards, these octant encodings are concatenated before the global encoder decreases the dimensionality further, as seen in Figure 6. The decoder mirrors the encoder and creates a reconstruction $\hat{y}_i$ for each y_i from the global latent space. Through this octant-wise encoding, the architecture incorporates knowledge about the LHC topography without using additional input features while keeping the network size small to avoid overfitting.

Layers. A separate Octant Encoder is used for each of the $n_c = 8$ octants as each octant i shows a different behaviour and has a different number of connected BLMs N_i. As seen in Figure 7a, the Octant Encoder consists of three layers and its latent space has $n_o = 8$ features. The latent spaces of all n_c octants are concatenated and passed as the input to the Global Encoder shown in Figure 7b, which has a latent space size of eight. After each layer, a Dropout of 0.1, batch normalization, and a ReLU activation function are used. The Octant Decoder and Global Decoder mirror this structure.

Data Normalization. The data from each BLM is normalised to the value range [0, 1] separately by dividing through the respective BLM's maximal observed beam loss measurement in the training set. Even though it loses information about the *Always-Low-Loss BLMs*' existence as all BLMs now have the same value range of [0, 1], it treats every input feature equally important.

Training. Training on a mix between Run-2 and Run-3 data would result in the model creating a cluster for each run instead of anomalies. Therefore, we use Run-2 data for training and Run-3 data for testing. Our proposed octant-wise autoencoder uses the *MSE-Loss* as a reconstruction loss:

$$L_{MSE} = \frac{1}{N} \cdot \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{7}$$

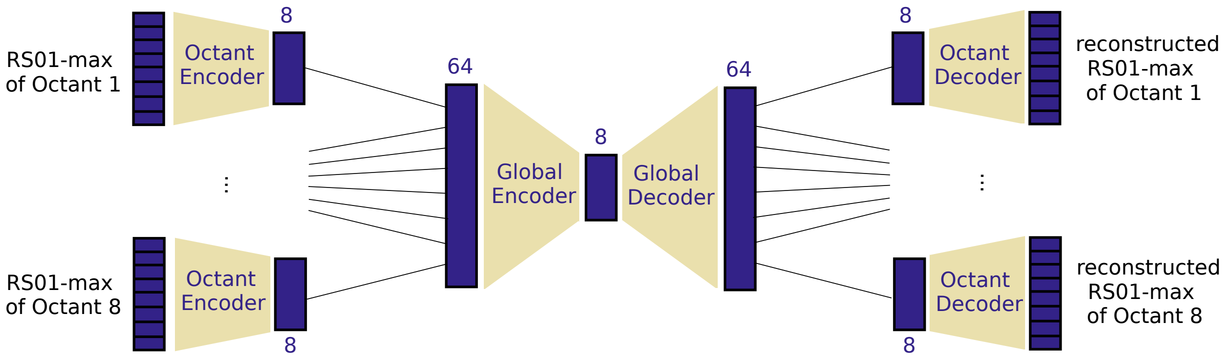


Figure 6: Our proposed octant-wise autoencoder for unsupervised anomaly detection. The octant-wise grouped features are encoded before being concatenated and encoded into a global latent space. The decoder mirrors this encoder.

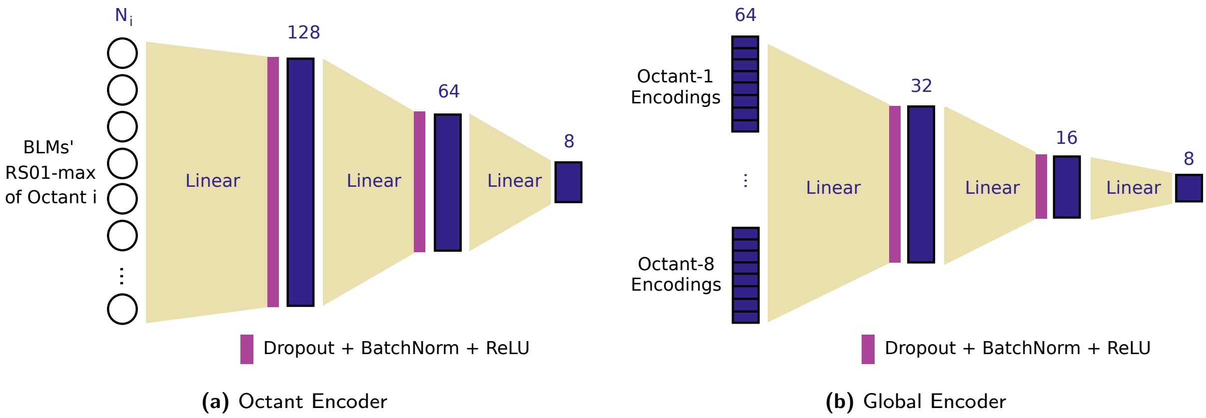


Figure 7: Layer composition of the proposed octant-wise autoencoder. A separate Octant Encoder for each octant encodes the RS01-max to a latent space of size 8. The Global Encoder further encodes the concatenated latent spaces of all Octant Encoders to a global encoding. Each later except the bottleneck use a dropout of 0.1 and batch normalization for more stable training.

6 Results

Despite the small dataset size, our proposed methods achieve in combination with one another an accuracy above 98% on average for ASD and UFO pattern detection. The sample size for other anomalies such as 10 Hz dumps is not sufficient to draw a confident conclusion.

6.1 ASD-Pattern Detection

Because ASDs have relatively high beam losses in specific regions, they are the easiest anomalies to spot within the data. Our method achieves very high accuracy for the final prediction even without information about the LHC topology.

Similarity Model. The training set consisted of all data from Run-2 and a few dozen samples randomly chosen from Run-3. The resulting 4-dim. similarity encodings of the test set form two clearly separated class clusters in all four of these feature dimensions. This is shown in Figure 8 where the left plot shows the first two feature dimensions of the similarity encodings, and the right plot shows the last two feature dimensions. A split between the two class clusters can be observed in latent feature 3 at $t_a = 33.47$ that separates the ASD encodings from other beam dump encodings. Beam dumps with an similarity encoding below t_a in dimension 3 are classified as ASDs. Using this threshold line t_a, the model achieves an accuracy of 99.5% for the test set. As seen in the confusion matrix shown in Figure 9, all ASD tests have been correctly classified and only one non-ASD sample has led to a False Positive. The False Positive is from a beam dump that showed unusual high loss readings due to a large perturbation from the electrical network (17-AUG-2022 18.01.39).

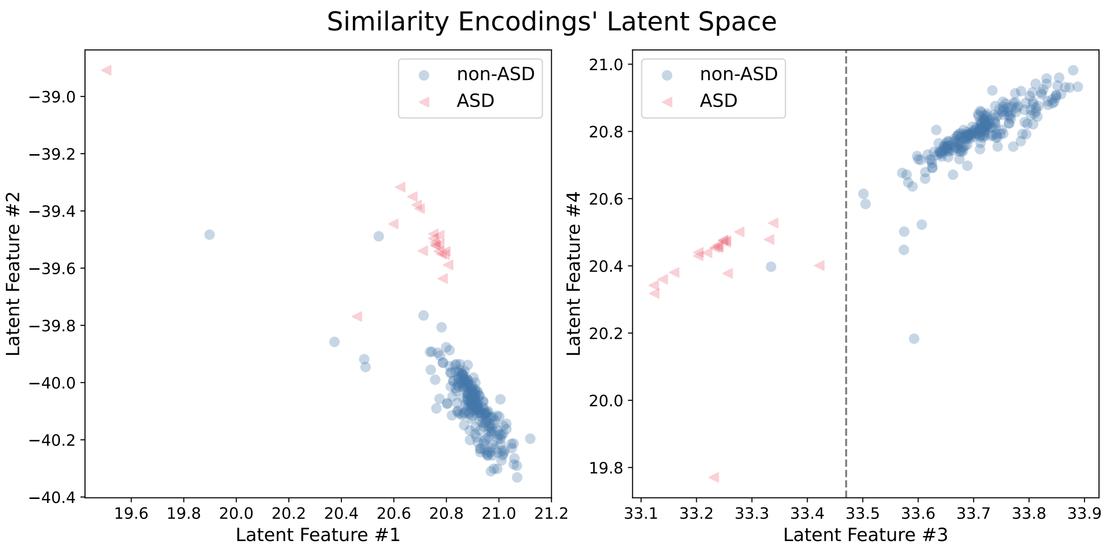


Figure 8: Similarity encodings of the ASD-Pattern Detection model. The similarity encodings of the test set over all four latent feature dimensions are shown. The similarity encodings from ASD tests form a clearly separated cluster from other events. A threshold line t_a is visualised in dimension-3 that is used for the classification of beam dump events. Beam dumps with a latent feature value smaller than t_s are classified as ASDs.

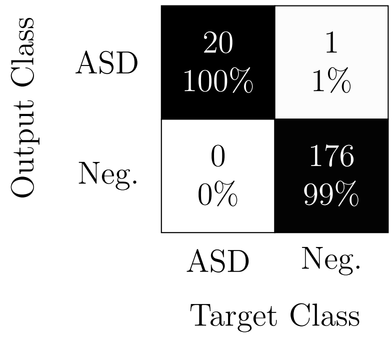


Figure 9: Confusion matrix of the ASD-Pattern Detection model. All ASD tests have been correctly classified and only one negative sample was incorrectly classified as an ASD. This False Positive has abnormally high loss readings in octant-2&3 because of an electrical network perturbation.

Octant-wise Autoencoder. Because ASDs have unusually high beam losses in Octant-6 and Octant-7, these losses are also very prominent in the encodings of these octants. However, high beam losses in Octant-7 can also stem from other types of anomalies like missing beam-beam kicks. By defining a suitable class threshold line, the encoding of Octant-6 can be used for ASD pattern detection as well, achieving the same accuracy of 99.5% as the supervised similarity model. The difference is that this approach has a False Negative instead of a False Positive.

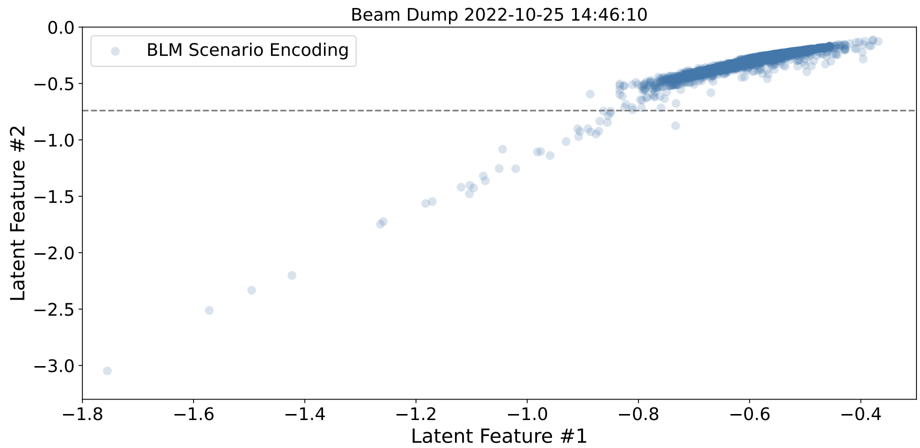
6.2 UFO-Pattern Detection

In order to classify UFOs among the beam dumps, we employ the *UFO-Losses Locally* model in combination with post-processing of predicted UFO votes. A beam dump is classified as a UFO dump if it has at least two votes $v = [v_1, v_2, v_3, v_4]^T$ with $v_2 < -0.75$ and passes the post-processing step. Examples of votes from two different beam dump events are shown in Figure 10. The standalone version of the model achieves an accuracy of 91.67% and in combination with the post-processing an accuracy of 98.33%. The resulting confusion matrix is shown in Figure 12. The one beam dump misclassified as non-UFO is a UFO in 13R6, which is more difficult to detect due to the intrinsic abort-gap losses in Octant-6.

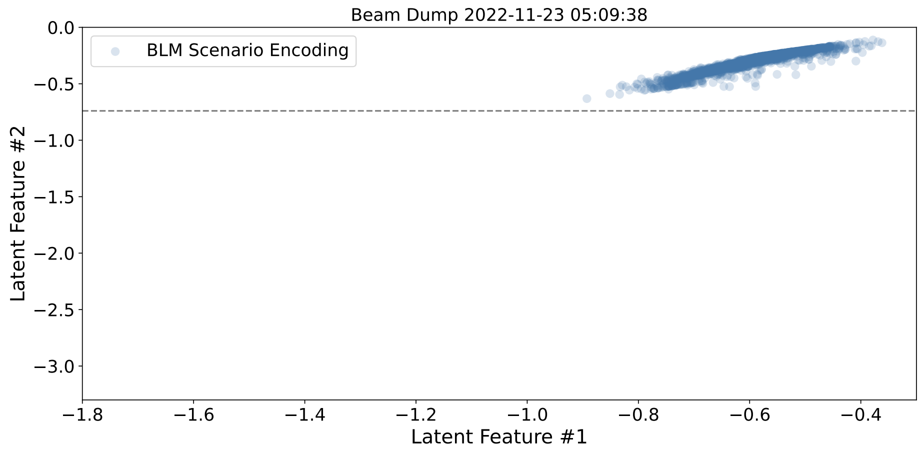
The test set consists of all beam dumps from Run-3. We excluded 96 non-UFO beam dumps from Run-3 as we use them for inference normalization. The network predicts higher absolute values (votes) for UFOs as seen in Figure 11 while non-UFO beam dumps form their own cluster. Depending on the chosen threshold line, the model achieves an accuracy of up to 91.67% with the confusion matrix shown in Figure 12 (a).

6.3 Abnormal Losses in Octant-7

The encodings from Octant-7 contain useful information for anomalous beam dumps while other octant encodings only show little or few anomalies, which is expected for some octants. Because of the collimators located in Octant-7, which clean the beams of stray particles, this region is more sensitive to beam losses. The encoding of Octant-7 learns a form of an abnormality scale within the observed octant (see Figure 13). This scale of severity seems to be similar to the standardization factor by the machine protection expert. However, this encoding only pertains to said octant and does not reflect abnormally high losses within other octants. The encodings from other octants and their global encoding are shown in Appendix A and are able to detect some anomalies in their respective region. With a threshold line at $t_s = 8.0$ within latent feature 1, as shown in Figure 13, the model achieves an accuracy of around 98.6% with the confusion matrix shown in Figure 14. Both abnormal beam dumps below the specified threshold t_s have no abnormally high losses in Octant-7 but instead in other octants. The false positive is from a beam dump with an orbit drift, resulting in higher beam losses within the observed octant.



(a)



(b)

Figure 10: Scenario encodings (votes) for each BLM for (a) a UFO event and (b) an anomalous beam dump with losses in Octant-3, 6, and 7. Each sample point represents the scenario encoding of a BLM before the global pooling operation. Because of the input normalization choice, test samples are classified via *min* pooling instead of *max*. Therefore, BLMs that observe UFO losses have a scenario encoding below the threshold line of 0.75 in latent feature #2. The smallest value within (a) is from BLMQI.02R8 and the UFO is at 01R8. The model not only detects the UFO but also locates its position within the ring.

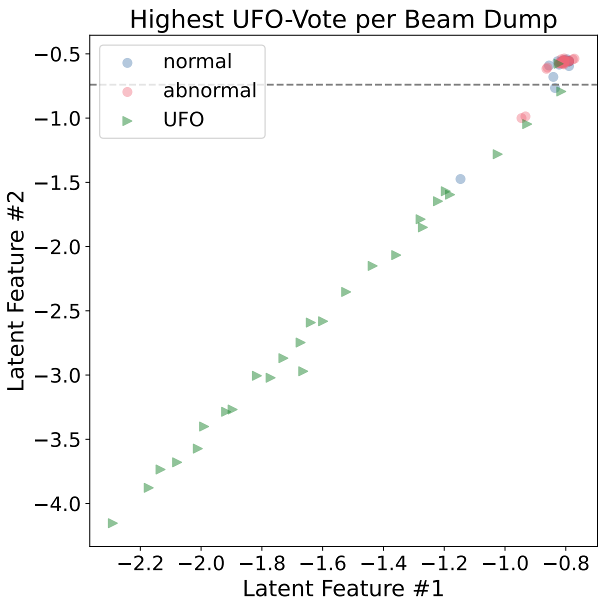
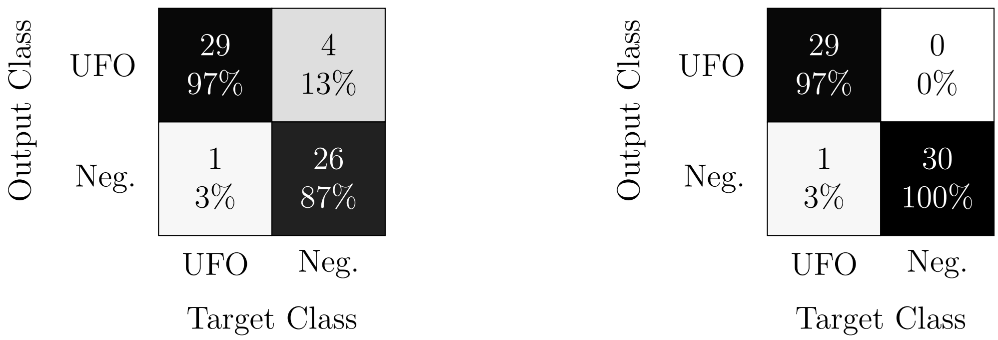


Figure 11: Scatter plot of highest UFO votes per beam dump. Each point depicts the smallest value of all BLM votes of a beam dump. Most UFOs have a BLM vote below the threshold line and most non-UFOs do not.



(a) **(b)**

Figure 12: Confusion matrices for the UFO-Losses Locally model (a) without post-processing, and (b) with post-processing. The post-processing invalidates UFO predictions if most of their votes above threshold stem from different octants, which is untypical for UFOs.

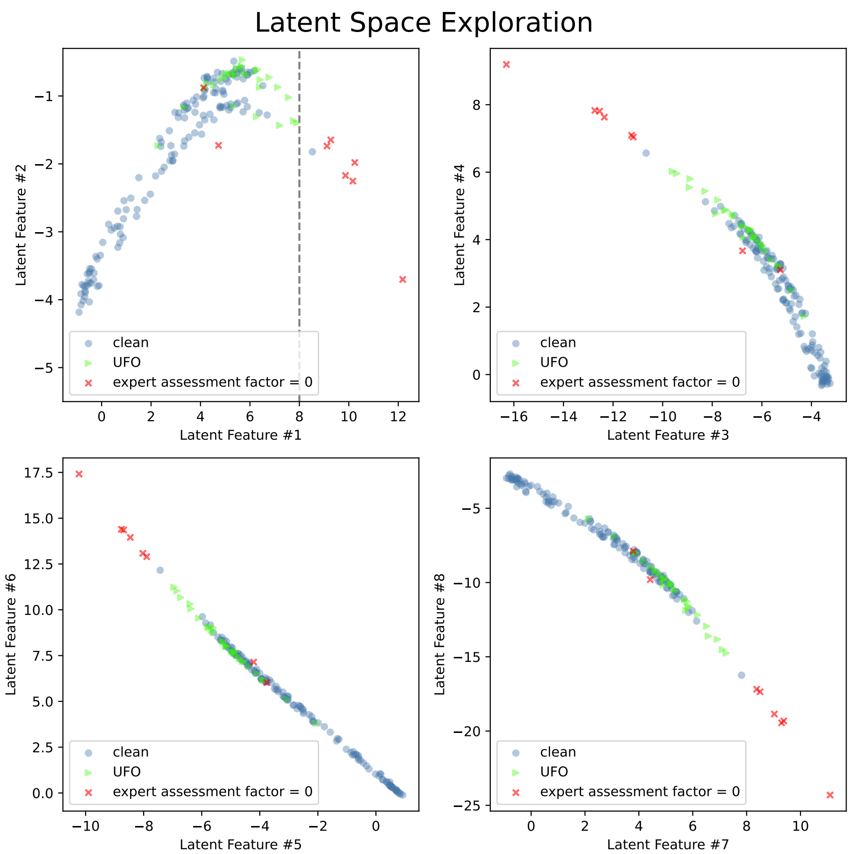


Figure 13: Latent space encoding of Octant-7. The red samples are beam dumps with an *expert assessment factor* of zero, meaning 'definitely anomalous'. Most of these beam dumps contain high losses within this octant, which leads to an encoding above the threshold t_s within dimension 1. This shows that features from this octant are very useful for detecting larger anomalies.

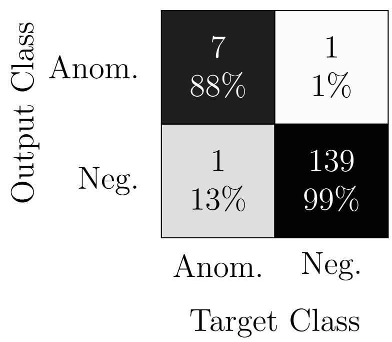


Figure 14: Confusion matrix of the Abnormal Losses in Octant-7 model. The threshold line for this classification is shown in Figure 13 Dimension-1. Both anomalous beam dumps below the threshold t_s do not show anomalously high losses within Octant-7 but in other octants. The False Positive contains higher than usual losses in Octant-7.

7 Ablation Study

This section covers the studies on why certain design and functional decisions have been made in the models. To be specific, it covers the choice of the pooling operation and the reason for the normalization in the UFO model.

7.1 Pooling Operation Choice

The *mean* global pooling operation is used for anomalies such as ASD tests that lead to high beam losses in many beam loss measurements. And the *max* global pooling operation is used for anomalies such as UFOs that only affect a few BLMs. As discussed in section 4, this choice of the global pooling operation within the supervised similarity model has a strong influence on its accuracy.

Multiple pooling operations. In the case of employing one global pooling operation across the BLMs within the model, we discard other possibly useful information about the data distribution from the BLM-scenario encoding. Therefore, a model using three different global pooling operations at once is tested. The three global pooling operations are the *mean*, *min*, and *max*, leading to $3 \cdot 4 = 12$ instead of 4 input nodes for the post-processing MLP. The resulting accuracy was similar to using the *mean* pooling operation. The reason is that the model predicted the same *min* values for every beam dump without deviation making their features obsolete — see Table 2. The *max* values only showed different features for a single dimension, and the *mean* operation was the only pooling operation that resulted in different values for all four feature dimensions.

Using *mean* for UFO Detection. Within the approach that uses multiple global pooling operations across the BLMs at once, the model led to more varying features for the *mean* pooling operation. But the resulting features from a *mean* pooling operation are less interpretable intermediate results than ones stemming from the *max* operation. While the intermediate results from the *max* operation can be post-processed with topological information of the LHC, the features from the *mean* operation cannot.

	Standard Deviation of BLM Vote's Features			
Pooling Operation	Dim. 1	Dim. 2	Dim. 3	Dim. 4
Min	0.0000	0.0000	0.0000	0.0000
Max	0.0000	**0.0500**	0.0000	0.0000
Mean [UFO]	0.0014	0.0009	0.0006	0.0006
Mean [non-UFO]	**0.0020**	0.0004	**0.0009**	**0.0008**

Table 2: Mean and standard deviation of scenario encodings after the global pooling operations. By using a model with multiple pooling operations, it learns a distribution in such a form that differentiates classes via the mean, resulting in the same *min* and *max* values for all beam dumps. The last two rows show the data distribution by the *mean* pooling operation grouped by class.

7.2 Data Normalization Choice

The choice of data normalization heavily influences the resulting similarity encodings of the UFO-Losses Locally model. We show that normalizing on clean beam dumps with a few beam dumps that contain abnormally high losses leads to better encodings. As seen in Figure 15 (a), using the same inference normalization as during training still clusters some beam dumps together with UFOs because these also have abnormally high losses in a small region. By normalizing on just clean beam dumps, a notion of abnormally high losses is enforced within the model. In combination with the learned BLM classes from the BLM-IDs, the model becomes more sensitive to lower values in the arcs where UFO signatures are usually present. By including 30 anomalous non-UFO beam dumps in the normalization calculation, the model can better differentiate between localized losses in the arcs and other low-loss regions.

False Positives. The False Positives are from non-neighbouring BLMs that are much further spread apart within the ring. This is a clear sign of not being a UFO because UFO-induced beam losses are clustered together in neighbouring BLMs. The reason why the model fails to learn this is that it has no prior information of BLM positions or BLM types. If the model had these data, the performance could be increased without post-processing. However, because the dataset is too small having more input information per sample complicates training.

False Negatives. The current votes are only collected from one of the four vote dimensions given by the BLM-scenario encoder. Some false negatives have votes with a slightly larger value in other dimensions, which are not used in the current approach. One false negative is a UFO in a region that usually observes higher losses, which makes it more difficult to spot UFOs in said region.

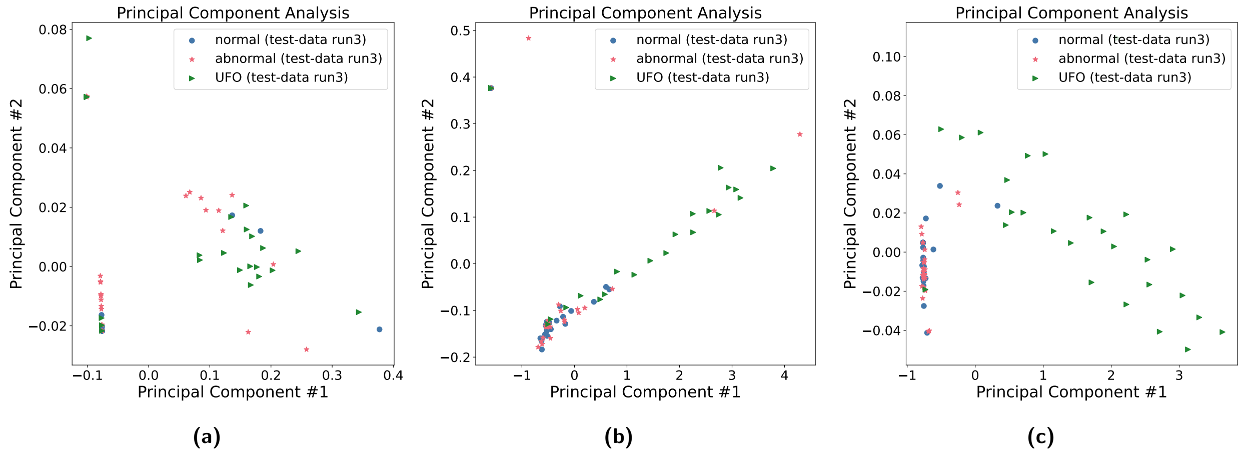


Figure 15: Principal Component Analysis of the UFO model's similarity encodings when using different inference normalization techniques. The test set data for the UFO Losses Locally model was (a) divided by 6 so that all BLMs have the value range [0, 2] that represents the entire value domain of BLM losses as during training, (b) normalized per BLM by dividing through the maximal loss seen in each BLM from only clean beam dumps, and (c) normalized per BLM by dividing through the maximal loss seen in each BLM from only clean beam dumps and a few anomalous beam dumps. Normalizing on certain beam dumps can improve the accuracy of the model.

7.3 Missing Beam-Beam Kick

The results from the Octant-7 encoding of the octant-wise autoencoder model show signs of assigning beam dumps with missing beam-beam kick (MBBK) events a much higher value in latent feature 1 — see Figure 17. The MBBK events are assigned a higher latent feature value because of their higher beam losses in said octant. The lower region of the right-side curve contains beam dumps with an MBBK beam-1 effect and the upper region contains beam dumps with an MBBK beam-2 effect. This indicates that the model can separate these kinds of events from other beam dumps. The model also shows hints of being able to differentiate between MBBK events stemming from beam-1 or beam-2.

7.4 10 Hz Events in Beam 1

The autoencoder model assigned 10 Hz events from beam 1 within the training set a similar encoding as an MBBK in beam 1. The latent space values of the Octant-7 Encoder are shown in Figure 18. Another difference between 10 Hz and MBBK events as seen by the model encodings is that the former are classified as ASD events (see Figure 16). The Octant-1 latent space encoding also assigned 10 Hz events a far different encoding than MBBK events, even though their encodings within Octant-7 are very similar. Via a combination of these models in the form of post-processing, it would be possible to label 10 Hz events. However, the dataset only contains two such beam dumps, which is not sufficient to form a conclusion on significant confidence levels.

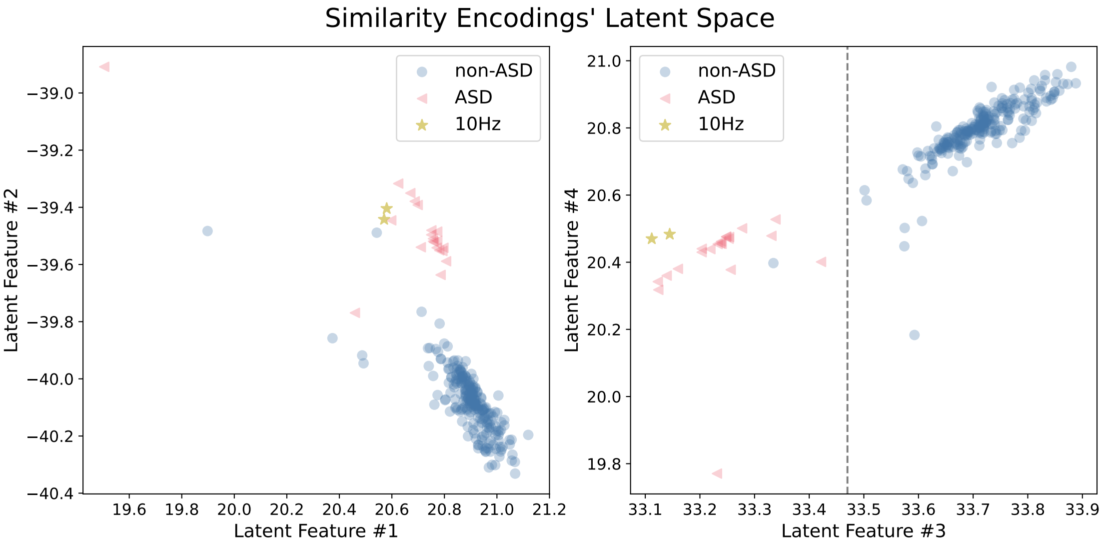


Figure 16: Similarity encodings from the ASD model for ASD, 10Hz, and 'other' beam dumps. Because 10 Hz events have high beam losses, the model incorrectly classifies them as ASD events too. However, with a combination of post-processing and the predictions from the other two models, these events can be separated.

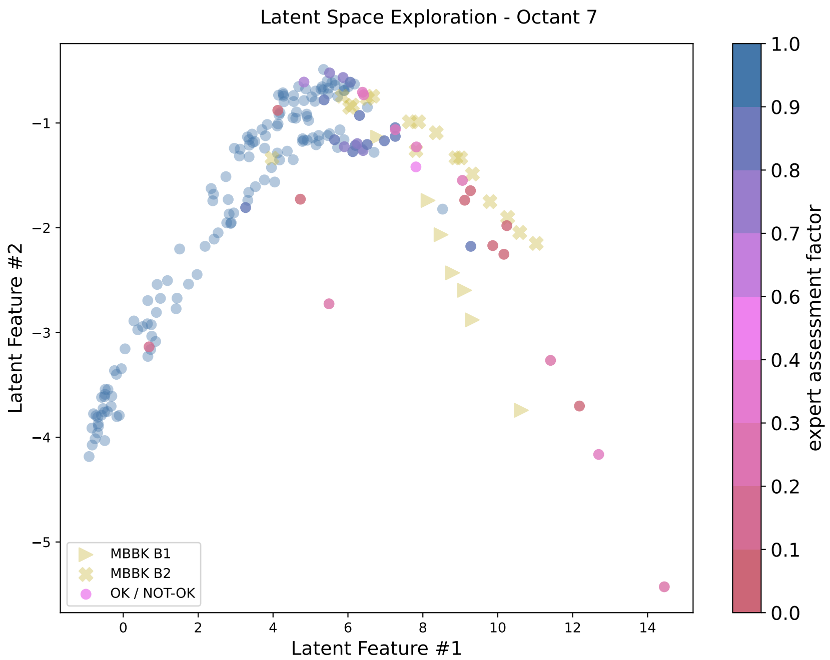


Figure 17: Octant-7 encoding in dimension-1 & 2 from the octant-wise autoencoder model. Samples in yellow are Missing Beam-Beam Kick (MBBK) events in Run-2. The MBBK events with an encoding at the bottom of the formed curve have an MBBK beam-1 effect while those with an encoding at the top of the curve have an MBBK beam-2 effect.

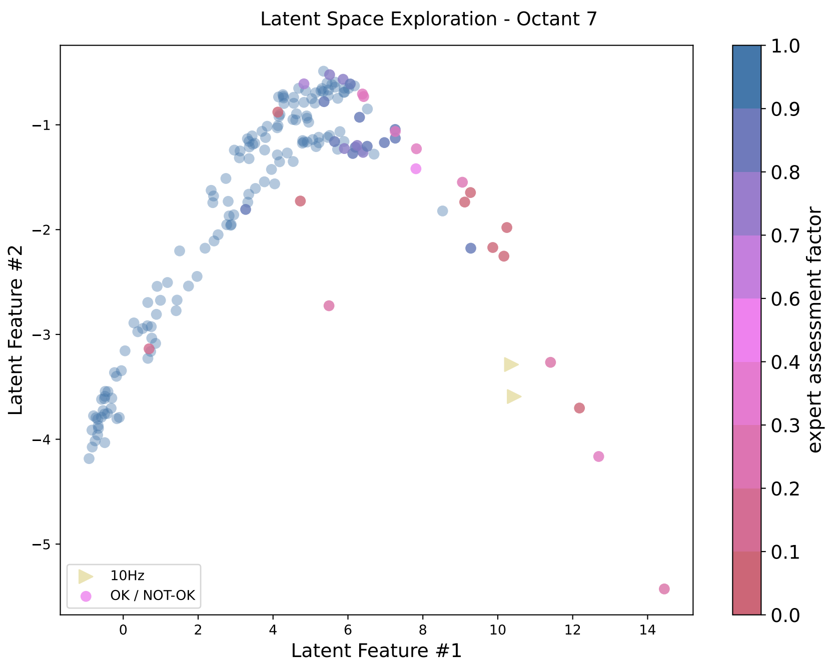


Figure 18: Octant-7 encoding in dimension-1 & 2 of the autoencoder model. Samples shown as yellow triangles are 10 Hz events in beam-1 from run2 (training set). These encodings are similar to MBBK beam-1 beam dumps within these dimensions of the Octant-7 latent space.

8 Conclusions

We have analysed the capability of neural networks for detecting and classifying anomalies from LHC beam losses and proposed three networks for classifying beam dumps into ASD, UFO, NOT-OK, and OK events. The networks are

- an equivariant model using similarity learning for ASD Pattern detection,
- an equivariant model for UFO localization using a voting system, and
- an autoencoder that creates features for each octant separately before encoding them further into global features.

In the case of ASD patterns, these beam dumps are more prominent within the data making them easier to spot. The model is therefore able to classify them very well even without the model having topology awareness of the LHC ring and their BLM positions.

For the PointNet-based UFO-Pattern model, the choice of inference normalization has a heavy influence on the predictions. Depending on this choice, the UFO pattern detection can be improved. We showed that normalizing on only clean beam dumps the model classifies beam dumps with high losses in *Always-Low-Loss BLMs* as UFOs. On the other hand, normalizing based on clean beam dumps and a few beam dumps with an abnormal expert assessment factor results in better accuracy for UFO classification. The intermediate results of the UFO model show high interpretability and are used in a voting technique to further improve the UFO classification accuracy. These UFO votes are being post-processed with topological information of the LHC. The BLM with the highest vote corresponds to the origin of the UFO.

The octant-wise autoencoder trained in an unsupervised setting is able to learn clusters of different meanings within each octant. The encodings of Octant-7 show an approximately linear scale of abnormally high losses within this octant, and the encodings of Octant-6 learn a similar scale of abnormally high losses. Meanwhile, Octant-2, -3, and -4 do not form a cluster for abnormal high beam losses, because these are usually not observed within these octants. However, our proposed network was able to identify one beam dump that unexpectedly had high loss readings within these octants. An extensive report on the octant encodings is given in Appendix A. Furthermore, the results indicate that our model would be able to classify missing beam-beam kicks and 10 Hz events with moderate post-processing of the octant encodings. But because too few data samples of these anomalies are available, we cannot form a conclusion on significant confidence levels.

9 Outlook

The proposed models can be improved by using LHC topology data such as the BLMs' position within the LHC ring and its type. Only ASD and UFO events are currently being classified. The automated detection and classification of other anomaly events such as 10 Hz events and missing beam-beam kicks is still an open topic.

Because of the small dataset size as well as the cost of labelling the data, Zero-Shot Learning and Few-Shot Learning for tabular data would be a fitting approach for classifying scarce events like 10 Hz anomalies. Our proposed neural networks already show that classifying such beam dumps with Deep Learning is possible, as seen in sections 7.3 and 7.4. A more sophisticated pre-processing step of the data could potentially improve the input information for the neural network. A possible method would be using the thresholds from *C. Wiesner et al.* [Wie+23] as statistical attention scores within the network. Additionally, using commonly-used positional encoding approaches like Fourier feature encodings to encode BLM identities to a lower dimension than one-hot encodings is still being considered.

The correlation between some of the latent features and the presence and degree of an anomaly might depend on various choices such as slight changes of the dataset, random neural-network weight initialization, and hyperparameters. To prevent that valuable information from getting distributed and hidden across several latent features, feature-disentanglement techniques such as dropout [Hin12] or β-VAE [Hig+17] can be used — in case they do not consider that information an entanglement of several factors of variation.

References

[Brü+04] Oliver Brüning et al. *LHC Design Report. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. doi: 10.5170.* Tech. rep. CERN-2004-003-V-1. URL: https://cds. cern. ch/record/782076, 2004.

[Eff+07] Ewald Effinger et al. "The LHC beam loss monitoring system's data acquisition card". In: (2007). DOI: `10.5170/CERN-2007-001.108`. URL: `https://cds.cern.ch/record/1027422`.

[Vig+12] William Vigano et al. *Comparison of three different concepts of high dynamic range and dependability optimised current measurement digitisers for beam loss systems*. Tech. rep. 2012.

[Hul22] Fredrik Hulphers. *Deep learning for anomaly detection in high-energy beam dump data from the Large Hadron Collider*. Tech. rep. 2022.

[Gas+16] Marek Gasior et al. "Introduction to beam instrumentation and diagnostics". In: *arXiv preprint arXiv:1601.04907* (2016), pp. 32–36.

[Wie+18] Christoph Wiesner et al. "Asynchronous beam dump tests at LHC". In: *9th Int. Particle Accelerator Conf.(IPAC'18), Vancouver, Canada*. 2018.

[Lec+22] Anton Lechner et al. “Dust-induced beam losses in the cryogenic arcs of the CERN Large Hadron Collider”. In: *Physical Review Accelerators and Beams* 25.4 (2022), p. 041001.

[Bél+22] Philippe Bélanger et al. “Charging mechanisms and orbital dynamics of charged dust grains in the LHC”. In: *Physical Review Accelerators and Beams* 25.10 (2022), p. 101001.

[Mir+19] Daniele Mirarchi et al. “Special losses during LHC run 2”. In: (2019).

[Lin+20] Björn Lindstrom et al. “Dynamics of the interaction of dust particles with the LHC beam”. In: *Physical Review Accelerators and Beams* 23.12 (2020), p. 124501.

[Wie+23] Christoph Wiesner et al. “Automated evaluation of LHC proton losses during high-energy beam dumps for the Post-Mortem System”. English. In: *Proc. IPAC’23* (Venice, Italy). IPAC’23 - 14th International Particle Accelerator Conference 14. JACoW Publishing, Geneva, Switzerland, May 2023, pp. 614–617. ISBN: 978-3-95450-231-8. DOI: `10.18429/JACoW-IPAC2023-MOPL038`. URL: `https://indico.jacow.org/event/41/contributions/1240`.

[Bau+21] Christoph Baur et al. “Autoencoders for unsupervised anomaly segmentation in brain MR images: a comparative study”. In: *Medical Image Analysis* 69 (2021), p. 101952.

[Ayt+18] Caglar Aytekin et al. “Clustering and unsupervised anomaly detection with l 2 normalized deep auto-encoder representations”. In: *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE. 2018, pp. 1–6.

[HA15] Elad Hoffer and Nir Ailon. “Deep metric learning using triplet network”. In: *Similarity-Based Pattern Recognition: Third International Workshop, SIMBAD 2015, Copenhagen, Denmark, October 12-14, 2015. Proceedings 3*. Springer. 2015, pp. 84–92.

[HRS11] Joe Hair, Christian M Ringle and Marko Sarstedt. “PLS-SEM: Indeed a silver bullet”. In: *Journal of Marketing theory and Practice* 19.2 (2011), pp. 139–152.

[Dev+18] Jacob Devlin et al. “Bert: Pre-training of deep bidirectional transformers for language understanding”. In: *arXiv preprint arXiv:1810.04805* (2018).

[WS09] Kilian Q Weinberger and Lawrence K Saul. “Distance metric learning for large margin nearest neighbor classification.” In: *Journal of machine learning research* 10.2 (2009).

[Hin12] Geoffrey Hinton. “Improving neural networks by preventing co-adaptation of feature detectors”. In: *arXiv preprint arXiv:1207.0580* (2012).

[Hig+17] Irina Higgins et al. “beta-vae: Learning basic visual concepts with a constrained variational framework.” In: *ICLR (Poster)* 3 (2017).

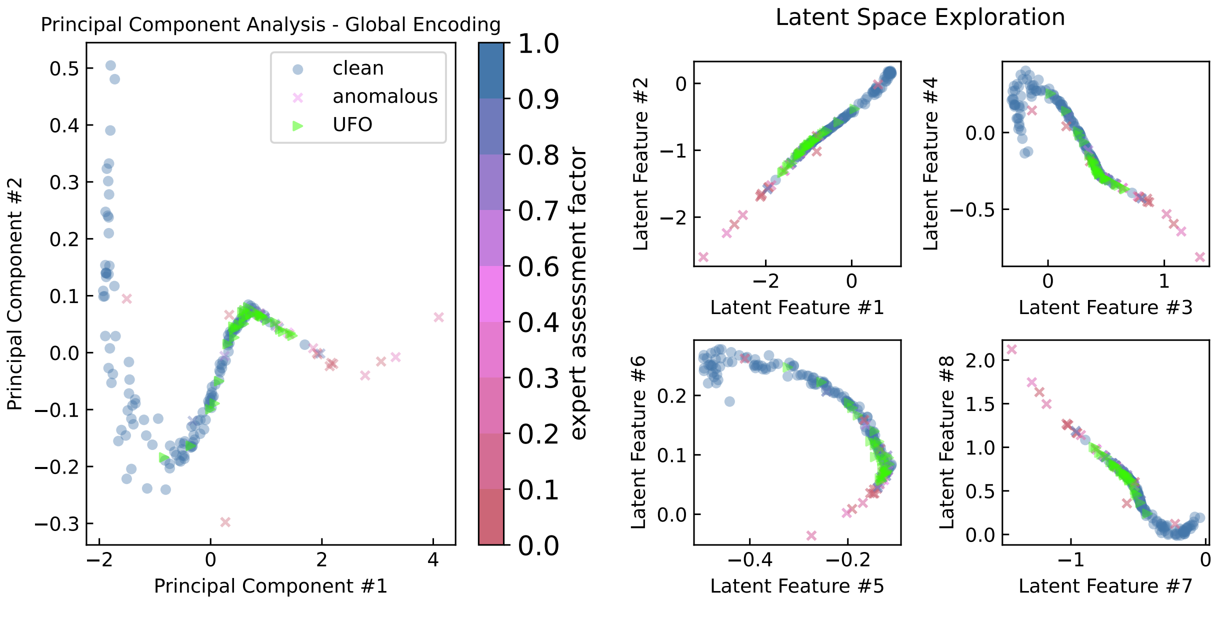
Appendices

A Octant Encodings of the Autoencoder

The following section covers the octant encodings as well as the global encoding from the octant-wise autoencoder. The sample points are colour-coded according to the expert assessment factor which is only present in the Run-3 data. A value of one corresponds to a nominal dump, also called a *clean* dump. A value of zero corresponds to a very anomalous beam dump event.

Global Encoding

As seen in Figure 19, the global encodings form a scale of how anomalous the beam losses seem, which is similar to what the Octant-7 encoding learned. Because most other octants usually do not observe many high losses, the global encoding is only influenced slightly of the other octant information. Table 3 lists selected beam dumps with additional comments that received a very high value within feature dimension 3 of the global encoding or appear to be out of distribution.



(a) (b)

Figure 19: The (a) Principal Component Analysis and (b) latent space dimensions of the global encoding from the autoencoder. The encoding seems to have learned a scale of how anomalous a beam dump is, which is similar to the encoding of Octant-7.

Timestamp	Dim.-3/4 Value	Additional Comments [1]
14-NOV-2022 09.42.13	(1.31, -0.81)	M1B1 RF trip in stable beams; Dumped due to RF trip M1B1 in SB (2460b). RMS H orbit of B1 moved by ∼100um. B2 dumped 155us after B1. Missing beam-beam kick signature with oscillations of ∼80um (B2 H) and ∼100um (B1 V) in last turn before B2 dump. Higher losses than usual in R7. AG pop of 2.2e9 (B1) and 1.4e9 (B1).
26-NOV-2022 17.38.24	(1.15, -0.64)	Programmed dump (2390b) for B1. B2 dumped ∼150us later due to BLMs. Missing Beam-Beam Kick signature with oscillations of +/- 80um (B2V), thus higher losses than usual in R7. Low AG population (∼3e8 for B1/2). Usual losses in IR6.
23-NOV-2022 06.09.38	(1.08, -0.59)	Dumped on losses in IR7 (TCLA.R7.B1) during ADT excitation of 180b during Quench Test MD (409b in B2, 13b in B1). Main losses in IR7, some losses in IR3/6.
27-OCT-2022 12.41.18	(1.02, -0.53)	Programmed dump. Abort gap losses + losses in IR7 (TCLA, TCSG and TCSPM). No significant orbit change.
23-NOV-2022 01.04.45	(0.87, -0.45)	Dumped on collimator temperature during slow ADT excitation of 180b during Quench Test MD (483b in B2, 13b in B1). Losses in IR6/7. RMS orbit movement of ∼25um before dump.
18-JUL-2022 02.28.45	(0.85, -0.45)	Losses got too high during off-momentum loss maps. Fill for Loss Maps with BBCW. Dumped on losses in IR7 (655ms RS) during +dp/p loss map. B1/2 H orbit change of ∼600 um due to off-momentum before dump. Usual losses in IR6/7 during dump.

Table 3: Additional Comments for selected beam dumps with a high encoding value in dimension 3 of the global encoding or those which appear out of distribution from the formed cluster. The table entries are sorted by the dimension 3 value and the encodings are visualized in Figure 19 (b). The beam dumps, which received a high encoding in dimension 3, show abnormally high losses within Octant-7 because of e.g. orbit drifts or missing beam-beam kicks.

Octant 1

As seen in the encodings of Octant-1 losses of beam dumps in Figure 20 and their comments in Table 4, the model is able to distinguish between clean beam dumps and some UFOs located within this region. UFOs within octant 1 receive an encoding further out from the formed cluster. By making use of a threshold of $t_1^1 = 0.55$ and $t_2^1 = 0.3$ in dimensions 1 and 2 respectively, eight beam dumps are outside the cluster and over the specified thresholds. All except one of these beam dumps are UFOs within octant 1. These beam dumps are listed in Table 4.

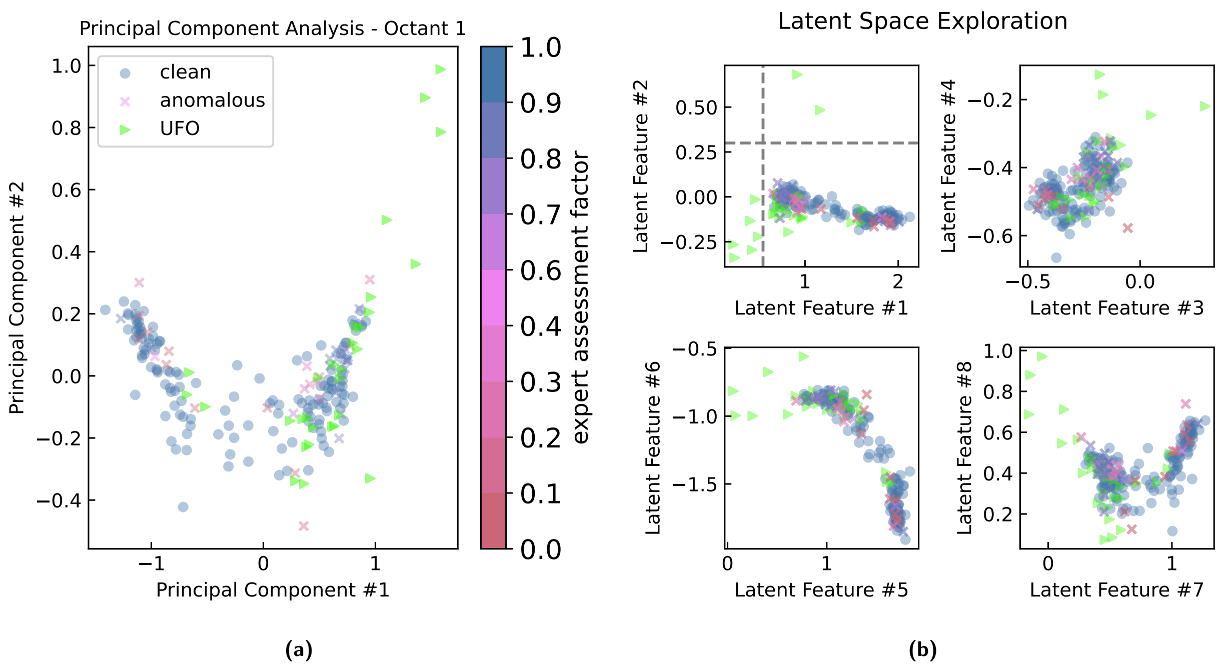


Figure 20: The (a) Principal Component Analysis and (b) latent space of the Octant-1 Encoding. Some beam dumps receive an encoding outside the formed cluster, which are mostly UFOs within Octant 1. Their comments are listed in Table 4.

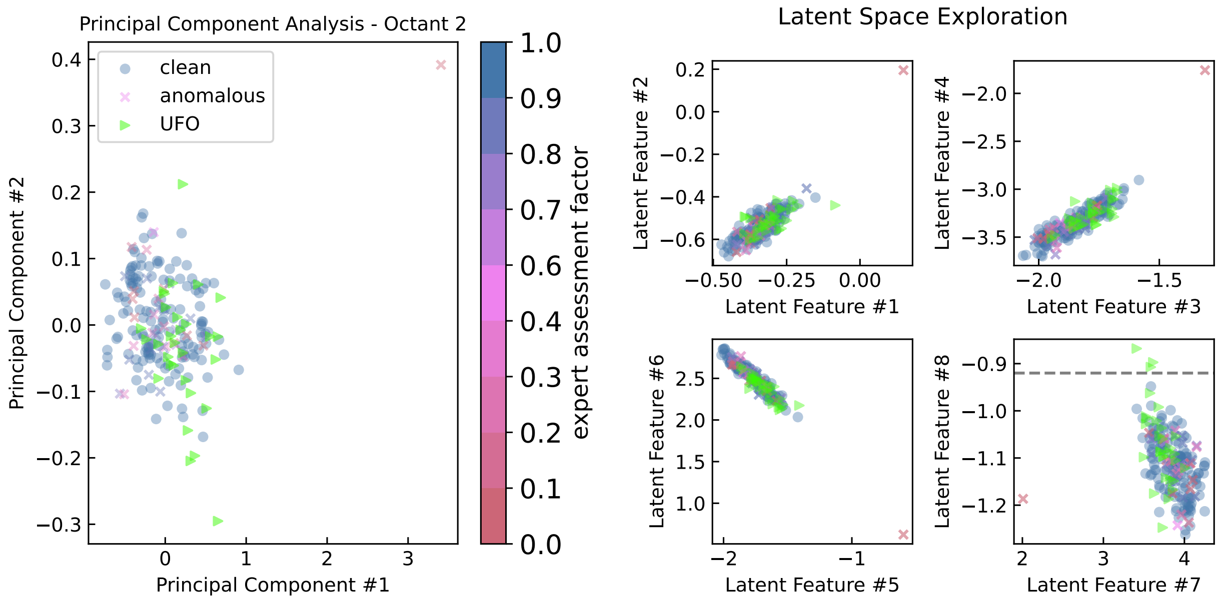
Timestamp	Additional Comments [1]
01-AUG-2022 04.31.42	UFO in 12R1
04-AUG-2022 23.13.55	UFO in 12R1
19-AUG-2022 05.46.24	UFO in 4L1
21-AUG-2022 19.48.23	UFO in 11R1
11-OCT-2022 05.22.47	UFO in Q10.R1
21-OCT-2022 01.26.25	UFO in 17R2
12-OCT-2022 00.30.55	UFO in 6R1
22-OCT-2022 16.02.47	UFO in MQ.6R1

Table 4: Additional Comments for selected beam dumps outside the formed cluster and over the specified thresholds t_1^1 and t_2^1. All except one of the beam dumps with encodings over the thresholds are UFOs within Octant-1. The other beam dump is an UFO within Octant 2.

[1] based on the Machine Protection Expert Comment and Operator Comment

Octant 2

The octant 2 usually does not observe abnormally high beam losses; however, our model's encoding shows three UFOs as well as a beam dump with unexpectedly high losses in this octant out of distribution (see Figure 21). The three UFOs are above the specified threshold of $t_8^2 = -0.92$ within dimension 8; their additional comments are listed in Table 5. The abnormal beam dump with a widely different encoding than other beam dumps is from 17-AUG-2022 18.01.39 with the mpe comment "Electrical perturbation, likely a lightning during thunderstorm triggered QPS in several RQ9 and RQ10 [...]". This beam dump shows abnormally high beam losses within Octant-2 to Octant-4, being far out of distribution from the formed cluster.



(a) (b)

Figure 21: The (a) Principal Component Analysis and (b) latent space of the Octant-2 Encoding. Because this octant usually does not observe high beam losses, the encoding forms a single cluster. Three UFOs lead to a higher encoding within dimension 8 of the latent space and one beam dump had unexpected high losses in this octant because of a lightning strike (17-AUG-2022 18.01.39).

Timestamp	Additional Comments [1]
10-JUL-2022 12.15.25	UFO in 17L2
19-JUL-2022 22.19.55	UFO in 17L2
21-OCT-2022 01.26.25	UFO in 17R2

Table 5: Additional Comments for the beam dumps over the threshold t_8^2 in dimension 8 as seen in Figure 21 (b). All beam dumps over this threshold line are UFOs within Octant-2.

Octant 3

Similar to Octant-2, Octant-3 usually does not observe high beam losses and only one cluster is being formed within its encoding. A few beam dumps with an encoding value below the threshold $t_7^3 = 2.05$ of dimension 7 of the latent space show various anomalies. But these anomalies do not show any pattern except having slightly higher beam losses within this octant. Same as in the Octant-2 encoding, one beam dump is far out of distribution, which is from 17-AUG-2022 18.01.39 again. The additional comments for beam dumps with an encoding below the specified threshold (as visualized in Figure 22) are listed in Table 6 except for the previously mentioned beam dump on 17-AUG-2022.

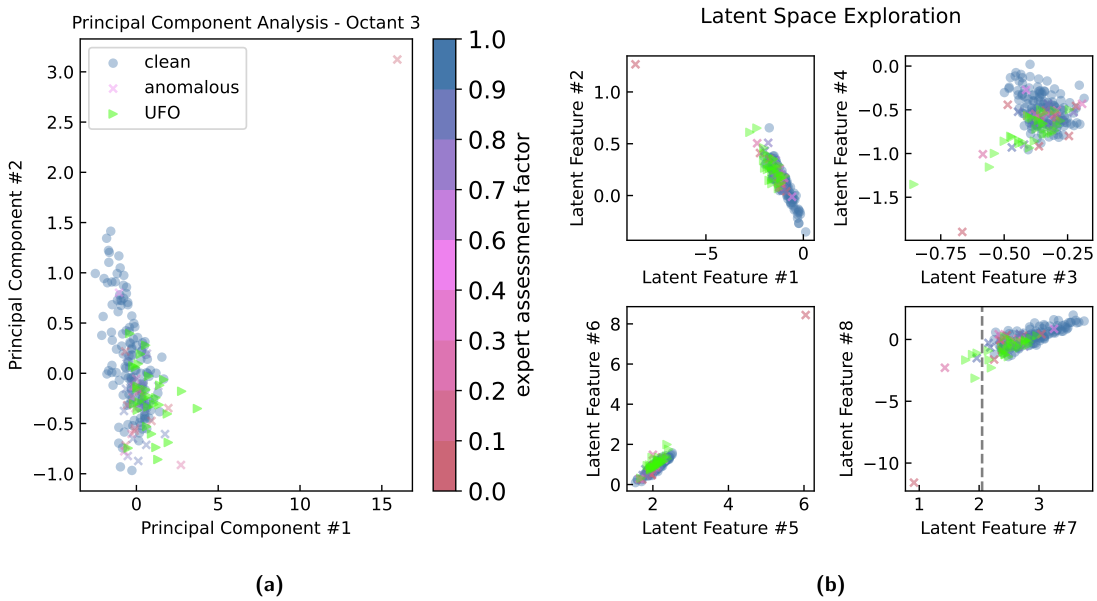


Figure 22: The (a) Principal Component Analysis and (b) latent space of the Octant-3 Encoding. Beam dumps below the threshold t_7^3 have slightly higher losses within this octant.

Timestamp	Additional Comments [1]
30-JUL-2022 04.03.39	UFO in 12R1
06-AUG-2022 06.30.47	UFO in 8R3
11-OCT-2022 05.22.47	UFO in Q10.R1
22-OCT-2022 16.02.47	UFO in MQ.6R1
02-NOV-2022 09.57.23	Programmed dump of 2460b physics fill. XPOC retrigger check B1 red (MKBs). Slightly higher losses than usual in IR6/7. Losses in R7 higher than in L7.
19-NOV-2022 13.14.19	Dumped on losses during loss maps for Pb ions at BLMQI.06R7.B1E10_MQTL (RS07) and TCLA.R7 (RS06). Orbit ok. Losses in IR6/7 and L3.

Table 6: Additional comments for beam dumps below the threshold t_7^3. They show slightly higher beam losses within Octant-3 with no other information.

Octant 4

As seen in Figure 23, the encodings of Octant-4 form two clusters of which the reason for the smaller cluster is a corrupted BLM card. Because of this, the BLMs BLMQI.22L4.B1I10_MQ and BLMQI.22L4.B1I30_MQ observe a sudden spike of beam losses of around 131 073 bits within their RS01. However, because our model only receives the RS01-max, it is not aware of the sudden spike and treats it as physically observed beam losses. Furthermore, two beam dumps are far out of distribution and have a large encoding value in dimensions 1 and 2 respectively. Their additional comments are listed within Table 7.

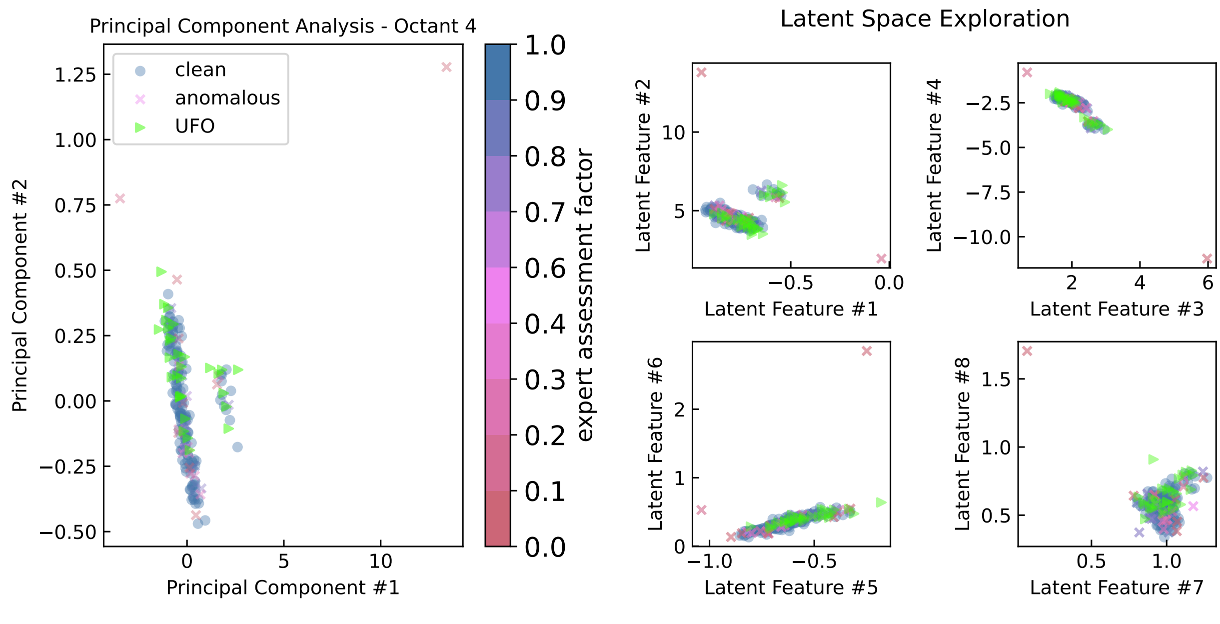


Figure 23: The (a) Principal Component Analysis and (b) latent space of the Octant-4 Encoding. A separate cluster is forming because of a corrupted BLM card leading to a sudden spike in the RS01. Moreover, two beam dumps are far out of distribution because of abnormally high beam losses.

Timestamp	Additional Comments [1]
17-AUG-2022 18.01.39	Electrical perturbation, likely a lightning during thunderstorm triggered QPS in several RQ9 and RQ10. Dumped in SB with 1935b (reduced #bunches due to SPS issue) due to electrical perturbation.
18-NOV-2022 13.59.41	Losses in IR4 during B2 wire scan. Dumped on losses with Pb ions at BLMQI.10L4.B2E10_MQML during wire scan (RS3/4). Orbit ok. Losses in IR4, IR6, R3. Fake BLM values (corrupted card) in R7.

Table 7: Additional comments for the two beam dumps with an encoding far out of distribution. These beam dumps have abnormally high values within dimensions 1 and 2 respectively. The one from 17-AUG-2022 also showed anomalies within the previous two octants.

Octant 5

Within dimension 1 of the Octant-5 encoding (see Figure 24) two clusters are forming, separating beam dumps with high CMS luminosity from those with a low luminosity. According to the split value of $s_1^5 = -0.22$ within dimension 1, the histogram of CMS luminosity grouped by these two clusters clearly shows that one cluster does not have high CMS luminosity, see Figure 25.

Furthermore, the model's encoding in dimension 4 shows values outside the formed cluster, which are UFOs and anomalies within this octant; their comments are listed in Table 8. One of these beam dumps is labeled as OK (no anomalies) but has multiple missing BLM values within Octant-5. Therefore, it also received an encoding outside of the formed cluster and being regarded as abnormal.

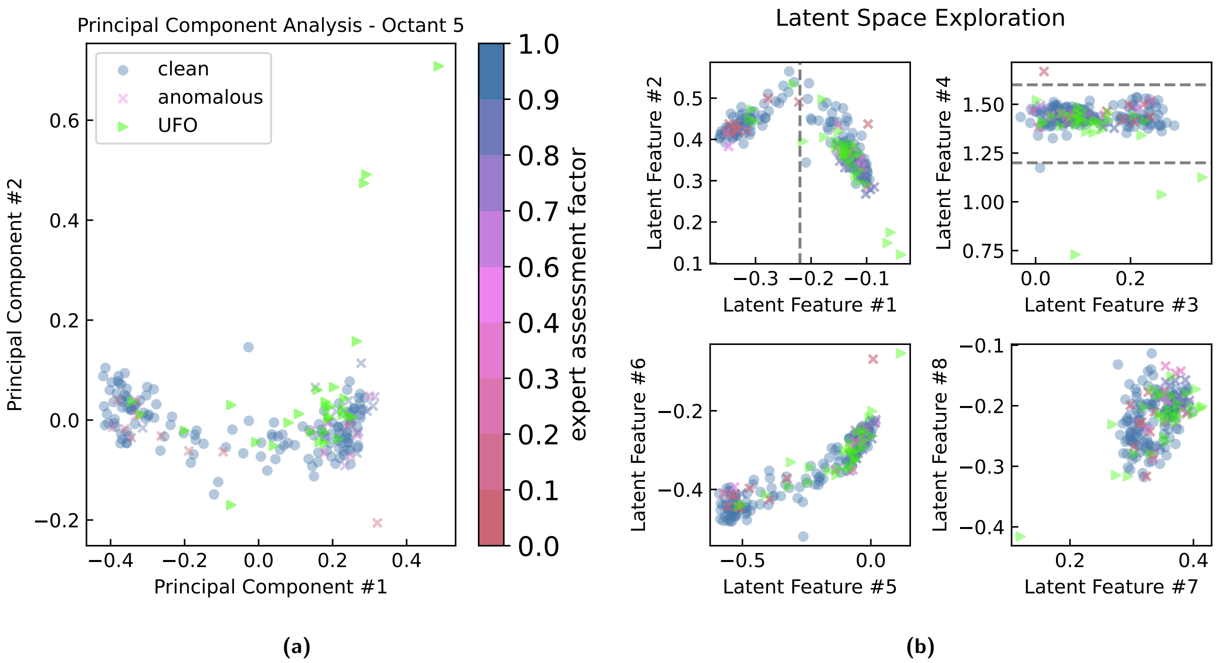


Figure 24: The (a) Principal Component Analysis and (b) latent space of the Octant-5 Encoding. The two clusters within dimension 1 are forming because of the CMS luminosity, and the few beam dumps outside the cluster within dimension 4 show abnormal beam losses within this octant.

Timestamp	Dim.-4 Value	Additional Comments [1]
17-AUG-2022 18.01.39	1.6684	Electrical perturbation, likely a lightning during thunderstorm triggered QPS in several RQ9 and RQ10.
28-OCT-2022 18.15.30	1.1742	A BLM crate tripped. Most likely due to failing temperature regulation in the rack. [Inspecting the dump again: multiple BLMs in octant 5 had missing values.]
08-AUG-2022 18.49.34	1.1251	UFO on TCL 6R5
09-AUG-2022 03.01.12	0.7291	UFO in 6L5
11-OCT-2022 19.09.28	1.0373	UFO in 6R5 (131%dump) (BLMTI.06R5.B1E10_TCL.6R5.B1)

Table 8: Additional Comments for beam dumps outside the defined thresholds of dimension 4. These beam dumps show abnormal losses within octant 5 or are UFOs within this octant.

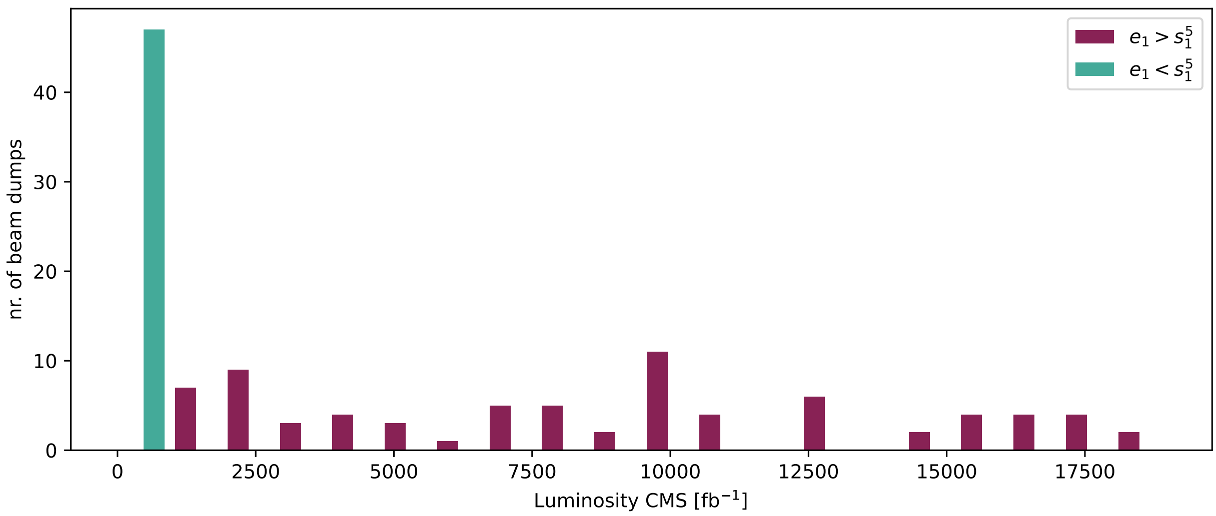


Figure 25: Histogram of CMS luminosities grouped by the two formed clusters within the Octant-Encoder's latent space e. Beam dumps with an encoding e_1 below s_1^5 in dimension 1 are shown in green, and beam dumps with an encoding e_1 above this value are shown in purple. We conclude that there is a correlation between the formed clusters and the CMS luminosity. This can be potentially attributed to the luminosity debris visible in this octant.

Octant 6

Similar to the previous octant, the encoding of Octant-6 is separated into two clusters with a high correlation to the beam intensity. Using a split value of $s_2^6 = 0.05$ to separate the two clusters as shown in Figure 26, one cluster consists mostly of beam dumps with low beam intensity while the other consists of dumps with higher beam intensities (see Figure 28).

Furthermore, the Octant-6 latent space encoded some beam dumps outside the formed cluster within dimension 3. Beam dumps below a defined threshold of $t_3^6 = -0.5$ all show anomalous behavior in their beam losses within Octant-6; their comments are listed within Table 9. One of these mentioned beam dumps has abnormally low losses within Octant-6 instead of high beam losses. Its analysis is given within Figure 27 showing unphysical values within Octant-7 and low beam losses in Octant-6.

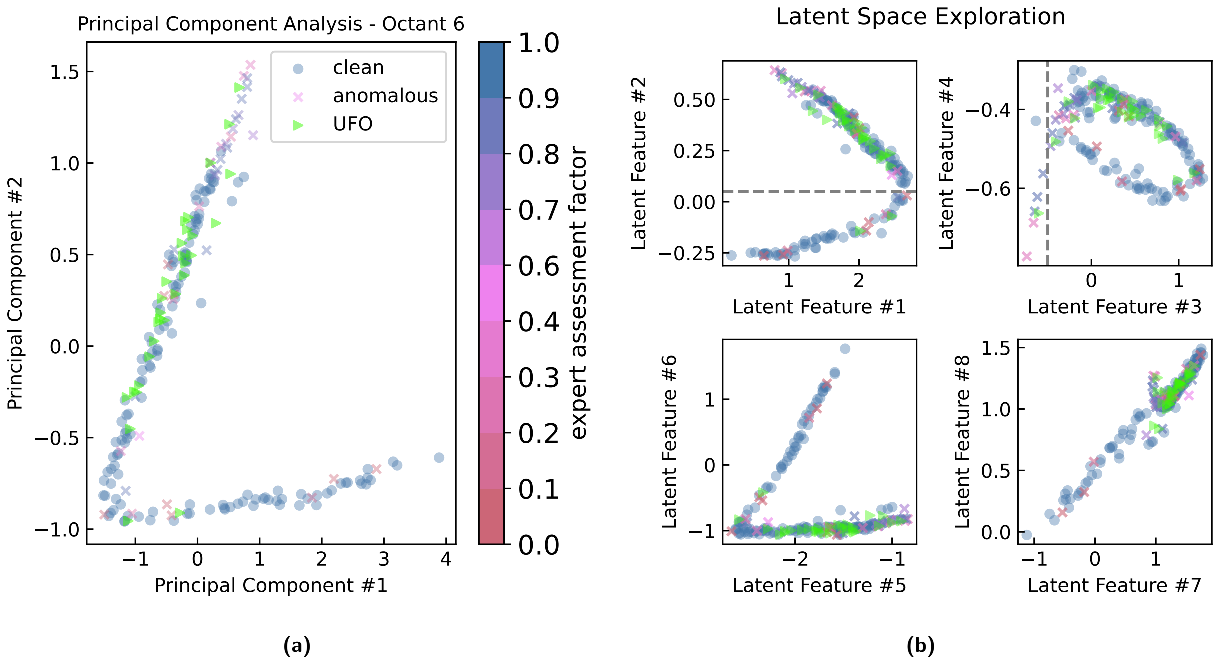


Figure 26: The (a) Principal Component Analysis and (b) latent space of the Octant-6 Encoding. The reason for the data separation within dimensions 1 & 2 is the beam intensity of the beam dumps. Beam dumps with an encoding above t_3^6 in dimension 3 show abnormal beam loss behavior within this octant.

Timestamp	Dim.-4 Value	Additional Comments [1]
09-AUG-2022 03.01.12	-0.67	The beam dump happened when we were moving the collimators just before moving in the roman pots. This is the second time it happens at exactly the same time. UFO: This time 6L5 when moving the TCL6 prior to the movement of the RP. No significant orbit change. Larger than usual losses in IR6: a TCSP and DFB BLM triggered (in XPOC all is good except the TCSP above warning).
30-SEP-2022 19.31.57	-0.69	Cryo Maintain S23 lost, controls problem; CM lost in S23. Larger than usual AG losses. No significant orbit change. PIC IPOC failed, PM injection and PM overall latched. "Required interlock" module red.
02-OCT-2022 18.00.31	-0.77	programmed dump, end of physics fill, after EoF tests (co-linearity and b6 knobs) in ADJUST; Programmed dump after EOF tests. Larger than usual AG losses. No significant orbit change. "Required interlock" module red."
02-NOV-2022 09.57.23	-0.66	Programmed dump of physics fill; Programmed dump of 2460b physics fill. XPOC retrigger check B1 red (MKBs). Slightly higher losses than usual in IR6/7. Losses in R7 higher than in L7.
05-NOV-2022 17.59.18	-0.56	End of physics fill. strange signals from 3 IR7 BLMs following the dump; Programmed dump of 2460b physics fill. No significant orbit change. Slightly higher losses in IR6. AG pop of 2.8e9 (B1) and 1.9e9 (B2). Wrong readings for some BLMs in R7 (card with corrupted memory).
16-NOV-2022 10.00.16	-0.62	Power converter fault(s); Physics fill (2400b) dumped by a trip of BBLR PC R1B2, most probably due to a short. BLMAI.05R6.B1E10_DFBLB above threshold.
19-NOV-2022 18.58.46	-0.43	Dump of a proton cycle validation fill with two indivs after the ion run; Programmed dump of validation fill after switching back to the proton cycle after the Pb ion test.

Table 9: Additional Comments for beam dumps below the defined threshold t_3^6 in dimension 3. They all show abnormal beam loss behavior within octant 6.

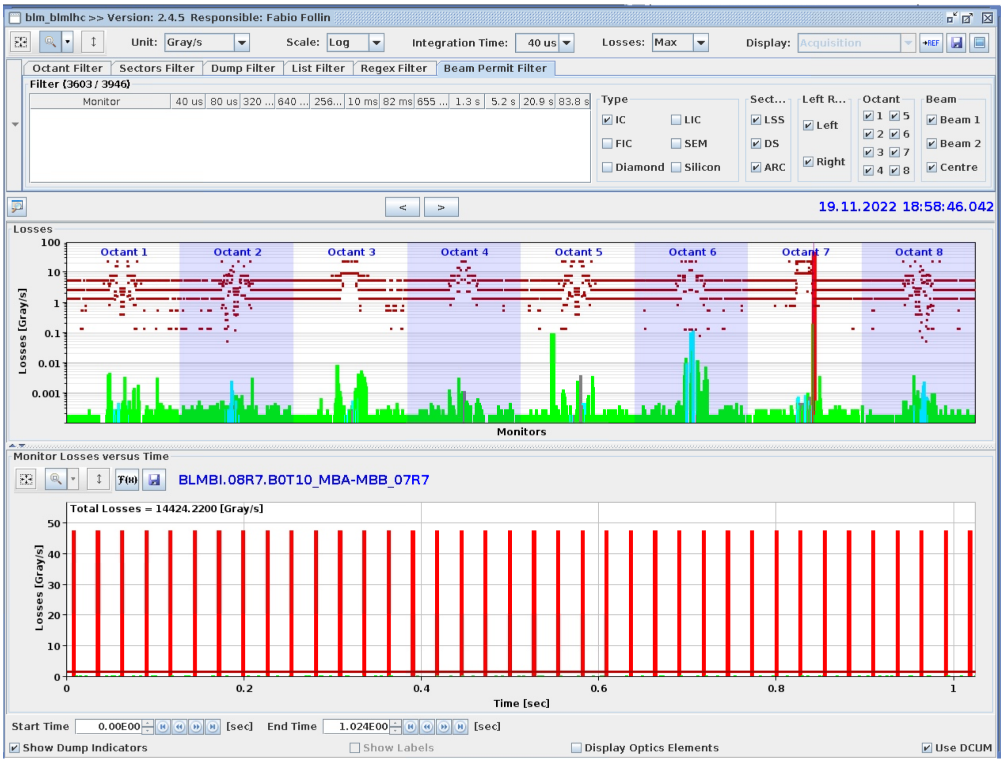


Figure 27: Analysis report from beam dump on 2022-11-19 at 18:58. The bottom row in red shows sudden and unphysical beam loss spikes within BLMBI.08R7. The middle row shows abnormally low losses within octant 6 and abnormally high losses within octant 7.

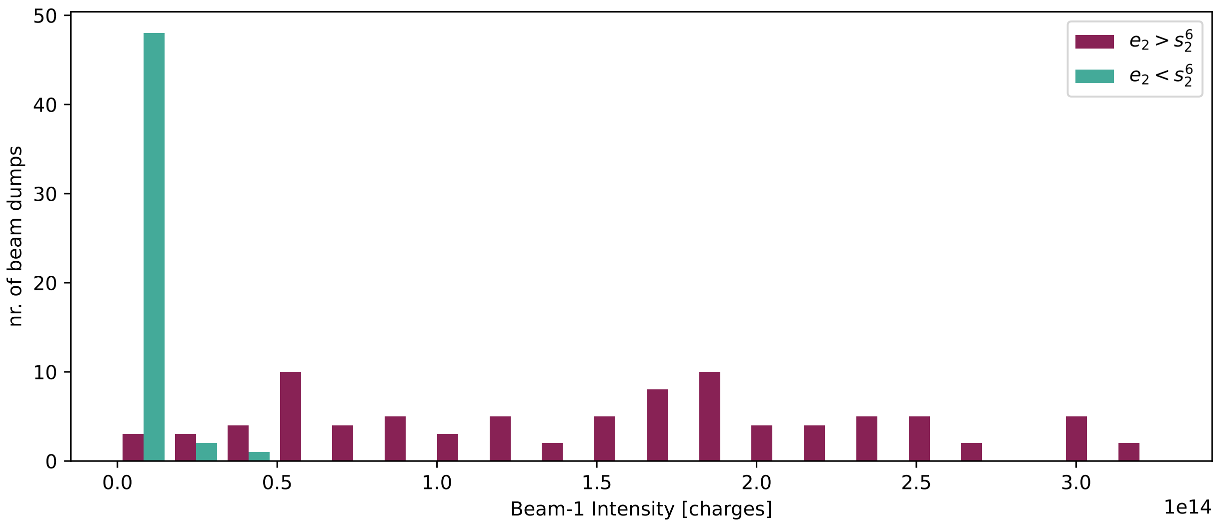
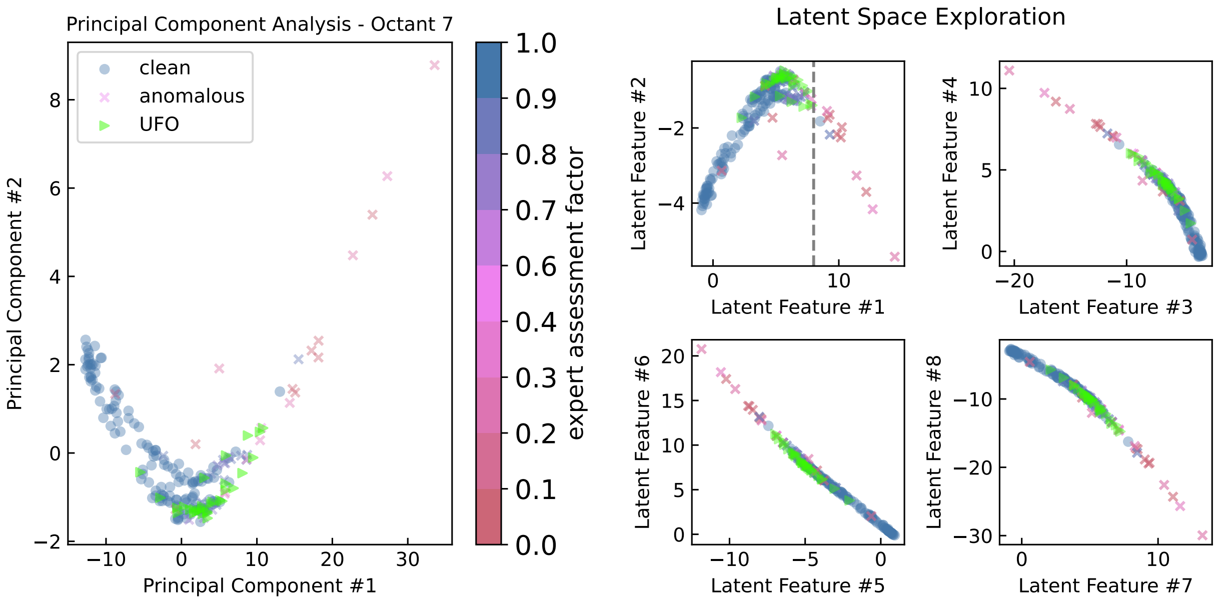


Figure 28: Histogram of beam 1's beam intensity grouped by the two clusters in the Octant-Encoder's encoding e in dimension 2 using the separation value s_2^6. Beam dumps with an encoding e_2 below this value in dimension 2 are shown in green, and beam dumps with an encoding e_2 above this value are shown in purple. We conclude that there is a correlation between the formed clusters and the beam intensity.

Octant 7

The encoding of Octant-7 correlates with abnormally high beam losses within this octant and, therefore, forms a scale of abnormality within said octant (see Figure 29). By using a threshold value of $t_1^7 = 8$, we can segregate the beam dumps with anomalous beam losses in this octant from other beam dumps. Comments for selected beam dumps are listed within Table 10 which are either out of distribution, above the threshold t_1^7, or labeled anomalous but are within the cluster of *clean* beam dumps. In the latter case, these beam dumps show anomalies within other octants but not within Octant-7. Therefore, these beam dumps are considered OK within Octant-7 by the model. One beam dump, which is labeled as OK, is above the previously defined threshold value and therefore labeled as anomalous by our approach. Upon further inspection, this beam dump contained unphysical beam loss values within Octant-7 as shown in Figure 30.



(a) (b)

Figure 29: The (a) Principal Component Analysis and (b) latent space of the Octant-7 Encoding. Beam dumps with an encoding above t_1^7 in dimension 1 contain anomalous losses within Octant-7. The few anomalous beam dumps below the threshold do not show anomalous losses in this octant but within other octants or their beam losses within this octant are not high enough.

Timestamp	Dim.-1/2 Value	Additional Comments [1]
18-NOV-2022 13.59.41	(3.26, -1.8)	losses in IR4 during B2 wire scan; Dumped on losses with Pb ions at BLMQI.10L4.B2E10_MQML during wire scan (RS3/4). Orbit ok. Losses in IR4, IR6, R3. Fake BLM values (corrupted card) in R7.
17-AUG-2022 18.01.39	(4.12, -0.87)	Electrical perturbation, likely a lightning during thunderstorm triggered QPS in several RQ9 and RQ10; Dumped in SB with 1935b (reduced #bunches due to SPS issue) due to electrical perturbation.
11-AUG-2022 17.27.58	(4.72, -1.71)	Beam lost during the cristal MD when exciting the beam with ADT with a gain too strong; Single bunch for crystla MD. Beam losses due to too high excitation from ADT. Losses in IR7 (BLMQI.06R7.B1E10_MQTL). XPOC B1 error (retriggering).
19-NOV-2022 13.14.19	(5.49, -2.72)	Collimation - Loss map; losses MQTL.6R7 during PbPb loss map (RS07); Dumped on losses during loss maps for Pb ions at BLMQI.06R7.B1E10_MQTL (RS07) and TCLA.R7 (RS06). Orbit ok. Losses in IR6/7 and L3.
19-OCT-2022 23.14.50	(8.5, -1.8)	BLM triggers in the long running sum in point 7; Dump by RS12 of TCP.B6R7.B2: see detailed analysis of the BLM issue with long RS bug in the code. Usual abort gap losses. Clean dump.
15-NOV-2022 23.18.30	(9.3, -2.2)	Power converter fault(s); Beams dumped in SB (2460b) after trip of RQT11.R8.B2 (caused by inconsistent current readings). AG population of 1.4e9 (B1) and 2e9 (B2), slightly higher losses than usual in IR6. Higher losses in R2 and L7 during the dump. Orbit changes < 0.5 mm

Table 10: Additional Comments for selected beam dumps. Some of the beam dumps below the threshold t_1^7 show abnormal losses within other octants are have lower beam losses compared to those above this threshold.

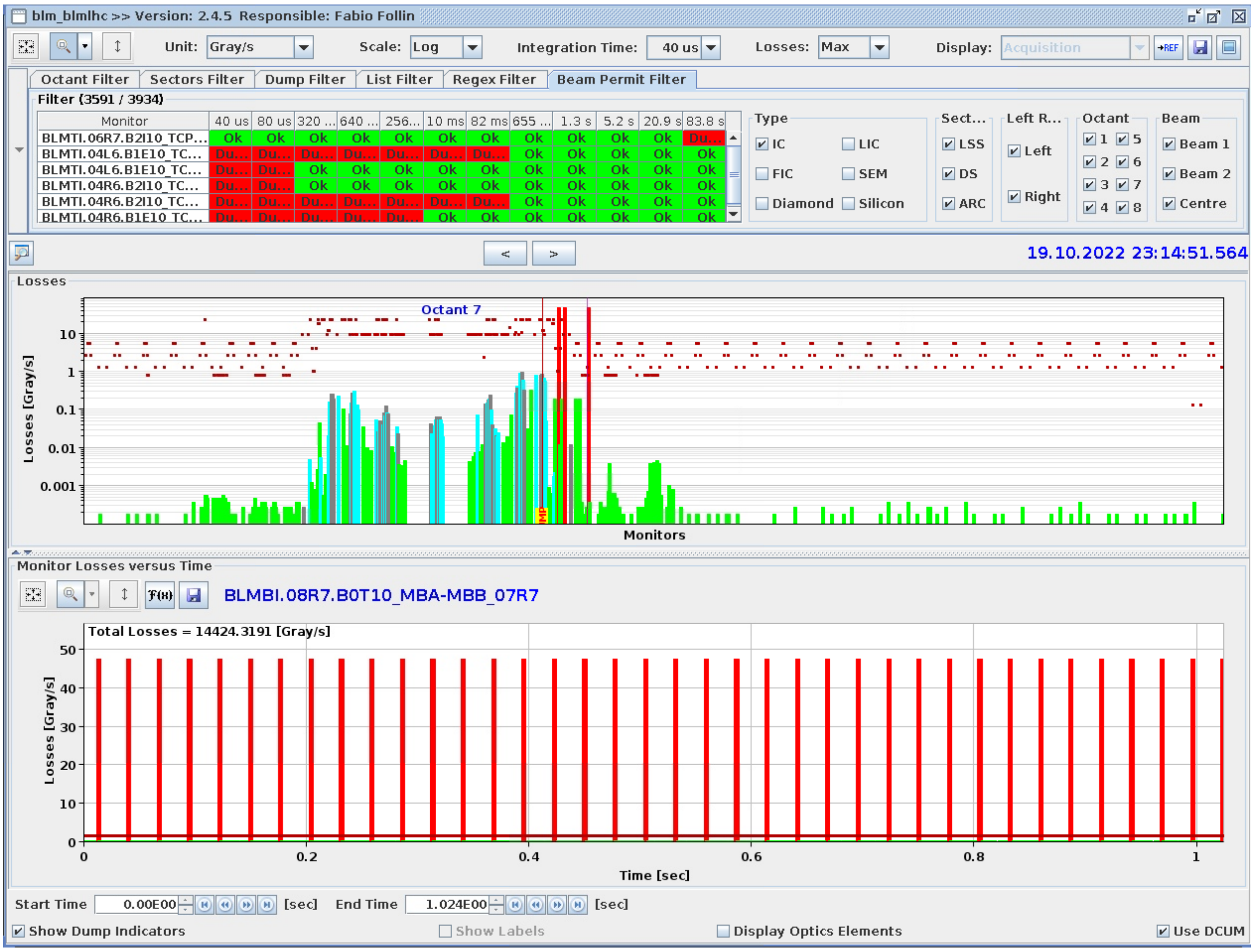


Figure 30: Analysis report from beam dump on 2022-10-19 at 23:14. The bottom row in red shows sudden and unphysical beam loss spikes within BLMBI.08R7. These (fake) high beam losses are also visible in the middle row as red columns.

Octant 8

Because Octant-8 usually does not observe high losses, the encoding is forming a single cluster with only three beam dumps outside it (see Figure 31). These beam dumps are above the thresholds $t_3^8 = -2$ and/or $t_4^8 = 1.2$ within dimensions 3 and 4 respectively, and their additional comments are listed within Table 11. Two of these beam dumps correspond to UFOs within this octant. The other beam dump contains unphysical beam losses within multiple BLMs of Octant-8 as seen in the detailed analysis view of Figure 32.

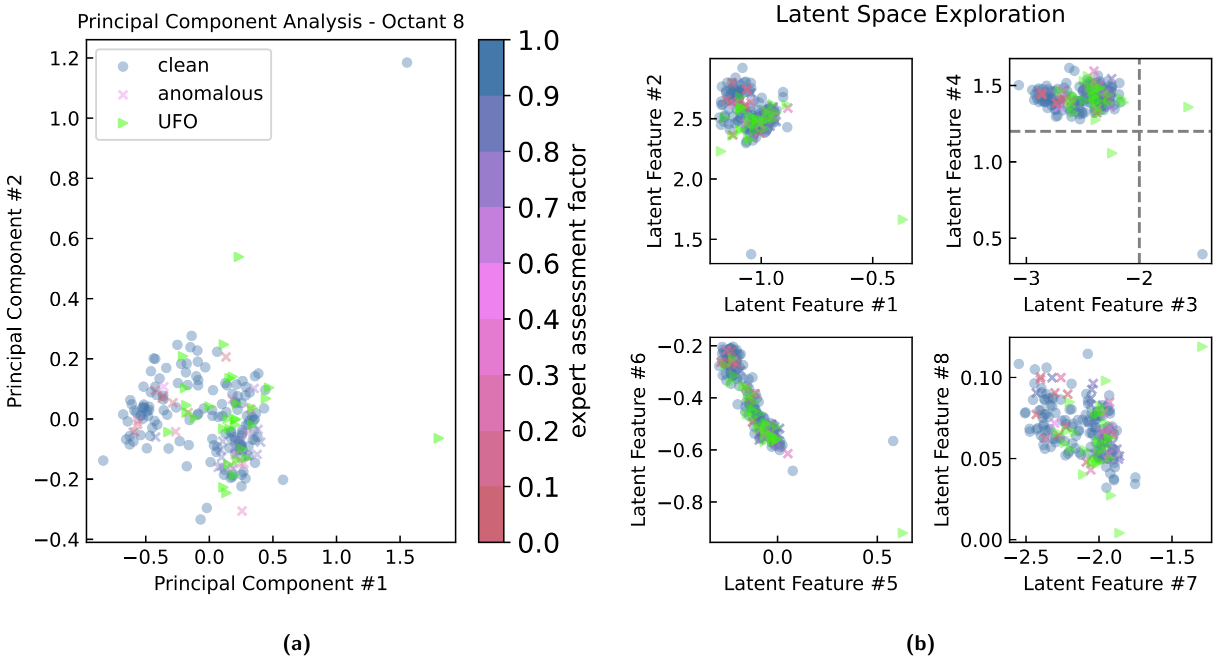


Figure 31: The (a) Principal Component Analysis and (b) latent space of the Octant-8 Encoding. Three beam dumps receive an encoding further outside the formed cluster. These beam dumps correspond to two UFOs within this octant and one beam dump with unphysically high beam losses.

Timestamp	Dim.-4 Value	Additional Comments [1]
28-JUL-2022 18.58.46	1.07	UFO in 27L8 when optimizing IPs (losses on BLMQI.27L8.B1E10_MQ). Good lifetime when going in collision following Q change.
25-OCT-2022 16.46.10	1.36	UFO 1R8; VeLo test insertion fill (1200b) dumped by UFO in 1R8 (first trigger by LHCb).
31-OCT-2022 10.01.35	0.4	programmed dump to give access to EN-EL for 48V power problem; Programmed dump of 2400b physics fill. Usual abort gap losses. Clean dump.

Table 11: Additional Comments for beam dumps with an encoding over the thresholds t_3^8 and t_4^8. All of them show anomalous high losses within octant 8.

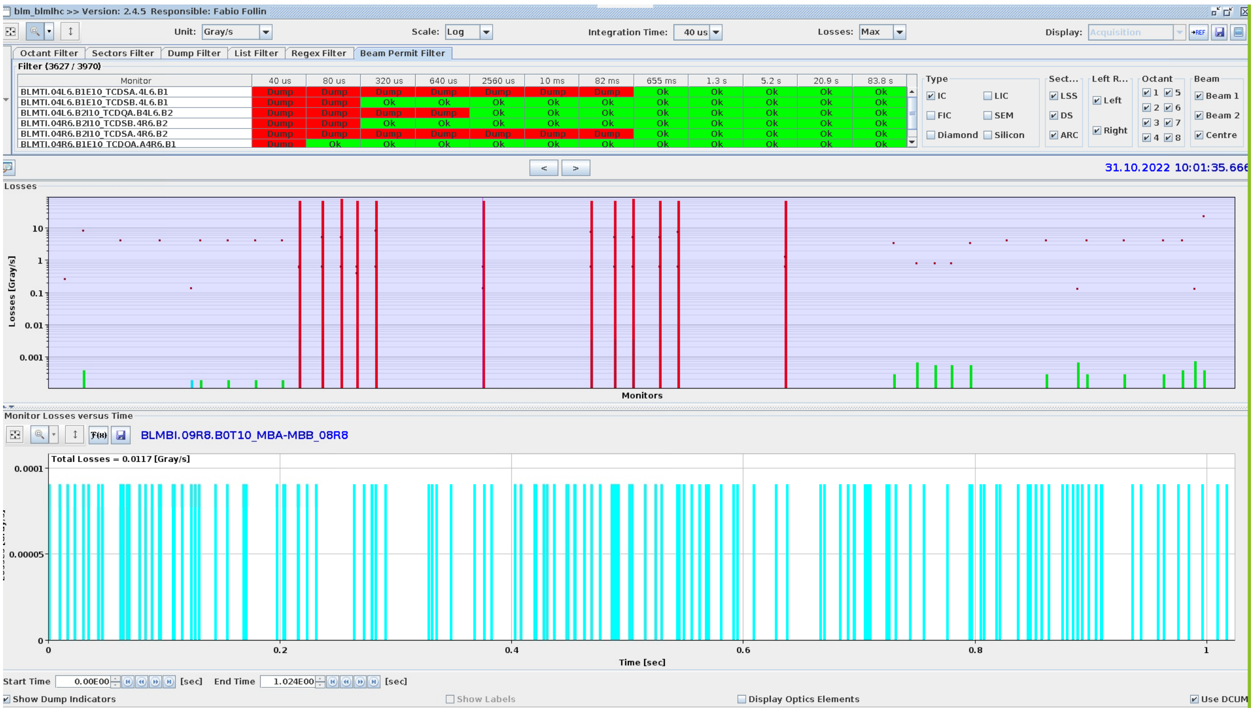


Figure 32: Analysis report on beam dump from 2022-10-31 at 10:01. Within the middle row, multiple columns representing beam losses within specific BLMs of Octant-8 are above the their thresholds and therefore marked in red. The detailed view of losses within one of these BLMs is given in the bottom row. The losses show abnormal behaviour by having multiple loss spikes that are unphysical.