



# Autoencoders for real-time SUEP detection

Simranjit Singh Chhibra<sup>1,2</sup>, Nadezda Chernyavskaya<sup>1</sup>, Benedikt Maier<sup>1,3,4,a</sup> , Maurizio Pierini<sup>1</sup>, Syed Hasan<sup>5,6</sup>

<sup>1</sup> European Organization for Nuclear Research (CERN), Meyrin, Switzerland

<sup>2</sup> Queen Mary University of London (QMUL), London, UK

<sup>3</sup> Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

<sup>4</sup> Imperial College, London, UK

<sup>5</sup> ETH Zurich, Zurich, Switzerland

<sup>6</sup> Scuola Normale Superiore (SNS) di Pisa, Pisa, Italy

Received: 25 November 2023 / Accepted: 26 January 2024

© The Author(s) 2024

**Abstract** Confining dark sectors with pseudo-conformal dynamics can produce Soft Unclustered Energy Patterns (SUEP), at the Large Hadron Collider: the production of dark quarks in proton–proton collisions leading to a dark shower and the high-multiplicity production of dark hadrons. The final experimental signature is spherically symmetric energy deposits by an anomalously large number of soft Standard Model particles with a transverse energy of  $O(100)$  MeV. Assuming Yukawa-like couplings of the scalar portal state, the dominant production mode is gluon fusion, and the dominant background comes from multi-jet QCD events. We have developed a deep learning-based Anomaly Detection technique to reject QCD jets and identify any anomalous signature, including SUEP, in real-time in the High-Level Trigger system of experiments like the Compact Muon Solenoid at the Large Hadron Collider. A deep convolutional neural autoencoder network has been trained using QCD events by taking transverse energy deposits in the inner tracker, electromagnetic calorimeter, and hadron calorimeter sub-detectors as 3-channel image data. Due to the sparse nature of the data, only  $\sim 0.5\%$  of the total  $\sim 300$  k image pixels have nonzero values. To tackle this challenge, a nonstandard loss function, the inverse of the so-called Dice Loss, is exploited. The trained autoencoder with learned spatial features of QCD jets can detect 40% of the SUEP events, with a QCD event mistagging rate as low as 2%. The model inference time has been measured using the Intel® Core™ i5-9600KF processor and found to be  $\sim 20$ ms, which perfectly satisfies the High-Level Trigger system's latency of  $O(10^2)$  ms. Given the virtue of the unsupervised learning of the autoencoders, the trained model can be applied to any new physics model that predicts an experimental signature anomalous to QCD jets.

## 1 Introduction

Hidden valleys, referring to hidden sectors with a confining gauge group, can produce dark quarks in proton–proton collisions at the Large Hadron Collider (LHC), leading to a dark shower and the production of a large number of dark hadrons ( $\phi_D$ ), analogous to QCD jets [1–3]. Depending on the details of the theory, the dark showers can follow large-angle emission, and dark hadrons do not arrange in narrow QCD-like jets. The decay of dark hadrons results in dark photons ( $Z_D$ ), which further decay to low-energy Standard Model particles with transverse energy ( $E_T$ ) of  $O(10^2)$  MeV, with the final experimental signature being high-multiplicity spherically symmetric Soft Unclustered Energy Patterns, or SUEPs. We focus on a well-motivated scenario, where SUEP is produced in exotic Higgs ( $H$ ) boson decays via gluon–gluon fusion and all dark hadrons decay promptly and exclusively to pions and leptons—an experimental nightmare scenario because of an overwhelming multi-jet QCD background.

We have developed an Anomaly Detection (AD) technique [4, 5] exploiting unsupervised deep learning for real-time SUEP detection in the High-Level Trigger (HLT) system [6] of the Compact Muon Solenoid (CMS) experiment [7] at the LHC.<sup>1</sup> Most AD techniques developed for high-energy physics applications rely on the online or offline reconstruction of collision events, which could be ineffective for reconstructing new physics signatures. We utilize raw quantities reconstructed in the HLT system, i.e., energy deposits in different sub-detectors (input for the online particle reconstruction). We have trained a deep convolutional neural autoencoder network (ConvAE) using QCD events, which is designed to detect events significantly different from the training data without prior knowledge of the signal characteristics. The data are three-channel images: two-dimensional (in the  $\eta$ - $\phi$  plane, where  $\eta$  is the pseudorapidity)  $E_T$  deposits in the inner tracker, electromagnetic calorimeter (ECAL), and hadron calorimeter (HCAL)

<sup>1</sup> The CMS trigger system is a two-tiered system. The first level (L1) triggers, implemented on custom electronics, process information from the calorimeters and muon system and select 100 kHz events of physics interest from 40 MHz collision events within a time interval of  $4\mu\text{s}$ . The second level (HLT) triggers run an optimized version of full-event reconstruction software on a farm of processors and further reduce the event rate to 1 kHz with a latency of  $O(10^2)$  ms.

<sup>a</sup> e-mail: [benedikt.maier@cern.ch](mailto:benedikt.maier@cern.ch) (corresponding author)

sub-detectors of CMS. Even though our work targets SUEP event tagging, the trained model could identify any other non-QCD-like signature, such as semi-visible jets [8] and emerging jets [9], spanning a large part of the landscape of the hidden valley. We plan on covering those new physics signals in our next paper.

This paper is structured as follows: Sect. 2 contains all the information about the QCD and SUEP event generation, simulation, and reconstruction. The event preselection is highlighted in the same section. The details about the ConvAE architecture and training are provided in Sect. 3. The trained ConvAE is applied to the QCD background and SUEP signal considering several Higgs boson mediator mass scenarios, and the performance results are given in Sect. 4. Section 5 concludes the paper.

## 2 Datasets

### 2.1 Generation, simulation, and reconstruction

Both multi-jet QCD and SUEP events are generated using PYTHIA 8.244 [10]. Detector effects are simulated and events are reconstructed with DELPHES 3.5 [11] using the CMS detector configuration. Proton–proton collisions at the LHC Run 3 center-of-mass energy of 13.6 TeV are considered, with approximately 50 collisions per bunch crossing (pileup). A pileup reduction is applied to the reconstructed tracks: the  $E_T$  deposits in the inner tracker by the pileup tracks are removed. No pileup reduction is applied for calorimeters, as per the HLT system’s configuration during Run 3 [6].

The model parameters for the QCD event generation are identical to those in [12]. For SUEP event generation, we used the SUEP\_Generator plugin [13], and the following parameter settings have been used:

- mediator masses ( $m_H$ ) = 125 GeV, 400 GeV, 700 GeV, and 1000 GeV
- dark meson mass ( $m_{\phi_D}$ ) = 2 GeV
- branching ratio of  $\phi_D \rightarrow 2Z_D = 100\%$
- dark photon mass ( $m_{Z_D}$ ) = 0.7 GeV
- branching ratios of  $Z_D \rightarrow \pi^+\pi^-, e^+e^-, \mu^+\mu^- = 70\%, 15\%, 15\%$
- dark temperature ( $T_D$ ) = 2 GeV [2]

Please note that we denote SUEP with a mediator mass of  $m_H$  GeV as SUEP( $m_H$  GeV) in the rest of the paper.

### 2.2 Event preselection

A preselection criterion is applied to both QCD and SUEP events based on the detector geometry and the typical characteristics of QCD events. The  $|\eta|$  of  $E_T$  deposits is required to be less than 2.5, as defined by the inner tracker coverage. The scalar sum of  $E_T$  of reconstructed electrons, muons, photons, and jets (denoted as  $H_T$ ) in an event is required to be  $>500$  GeV, which is a typical  $H_T$  requirement at the HLT after L1 trigger selection during Run 3. The event preselection efficiency is found to be  $\sim 85\%$  for QCD and  $\sim 2.1\%$ ,  $\sim 10.4\%$ ,  $\sim 18.4\%$ , and  $\sim 22.7\%$  for SUEP(125 GeV), SUEP(400 GeV), SUEP(700 GeV), and SUEP(1000 GeV), respectively.

The plots showing the  $E_T$  deposits for a QCD event and a SUEP(1000 GeV) event, which satisfy the preselection criterion, in the inner tracker, ECAL, and HCAL are given in Figs. 1 and 2, respectively. The ECAL is 25 times more granular ( $0.0174 \times 0.0174$  rad<sup>2</sup> in  $\eta$ - $\phi$ ) than the HCAL ( $0.087 \times 0.087$  rad<sup>2</sup> in  $\eta$ - $\phi$ ). Hence, each HCAL pixel is divided into 25 equal pixels to match the ECAL granularity. The final shape of the images is (288, 360, 3). For this particular QCD event, we can appreciate the typical signature of two jets in the  $\eta$ - $\phi$  plane. Also, we can acknowledge the sparse nature of the data—only  $\sim 15k$  pixels out of a total of 3,11,040 image pixels have nonzero values. This observation is valid in general for all QCD events. On the other hand, for the SUEP(1000 GeV) event, there is a clear spherically symmetric signature of a large number of low- $E_T$  particles in  $\phi$ . We normalized the  $E_T$  values such that they are of  $\mathcal{O}(1)$ , which facilitates more stable and faster ConvAE training.

## 3 Autoencoders

Our ultimate target is to reject QCD jets and identify anomalous signatures. The unsupervised learning nature of autoencoders (AEs) perfectly serves the purpose [14]. In its simplest form, an AE is a neural network (NN) that maps the input (ground truth) to a latent compressed representation (encoding), the so-called bottleneck, and then back to the input dimensionality (decoding), providing the output (reconstruction) as an approximation of the input. The quality of the AE training is measured in terms of a distance metric, the so-called reconstruction loss; e.g., for images, the loss function can be defined as the pixel-wise summed mean-squared difference between input and output. Once the AE is trained using background events with the objective of minimizing the reconstruction loss, the model is expected to poorly reconstruct anomalous signatures and result in large reconstruction loss values. A threshold on the reconstruction loss can be imposed to tag anomalies.

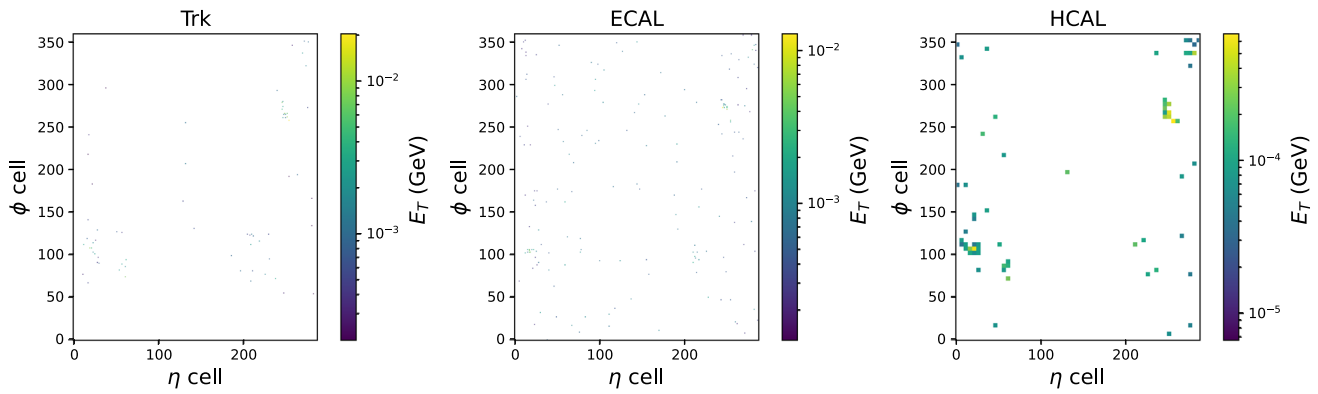


Fig. 1  $E_T$  deposits for a QCD event in the inner tracker (left), ECAL (middle), and HCAL (right)

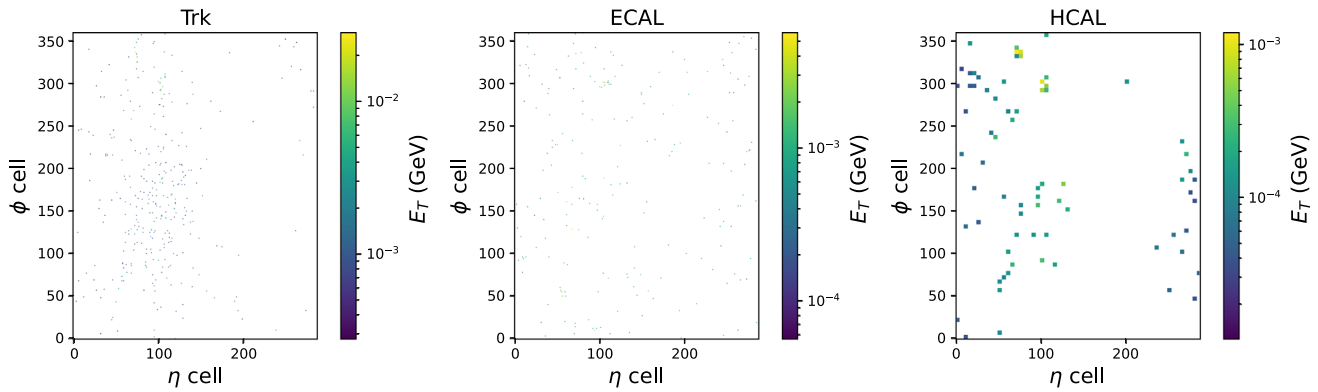


Fig. 2  $E_T$  deposits for a SUEP(1000 GeV) event in the inner tracker (left), ECAL (middle), and HCAL (right)

### 3.1 Architecture and training

The ConvAE architecture is detailed in Table 1. For encoding, there are five two-dimensional Convolutional NN (CNN) [15] layers with a `kernel_size` of (3, 3) and `strides` of (3, 3), (2, 2), or (2, 3). Up to the bottleneck, the input shape of (288, 360, 3) compresses to (6, 5, 8). The decoder part, which is exactly the mirror image of the encoder part, is composed of five two-dimensional Transposed CNN [16] layers with the same `kernel_size` and exactly opposite `strides`. The `padding` is set to 'same'. After each (Transposed) CNN layer, a `BatchNormalization` layer is introduced, which facilitates a more stable and faster training by bringing the mean of the output close to zero and the output standard deviation close to unity. The Parametric Rectified Linear Unit (PReLU) and Rectified Linear Unit (ReLU) [17] activation functions are used for all the hidden layers and the output layer, respectively. This model architecture results in a total number of trainable hyperparameters equal to 3,574,065. The model was implemented in Keras/TensorFlow [18, 19], and it was trained for 100 epochs (with early stopping enabled) with a `batch_size` of 128. The learning rate was set to  $10^{-2}$ , and the optimizer found to yield the best performance was Adam [20]. The training was performed using the NVIDIA® Tesla® V100 PCIe 32 GB Graphics Processing Unit [21], and ~25 min per epoch were consumed while model training.

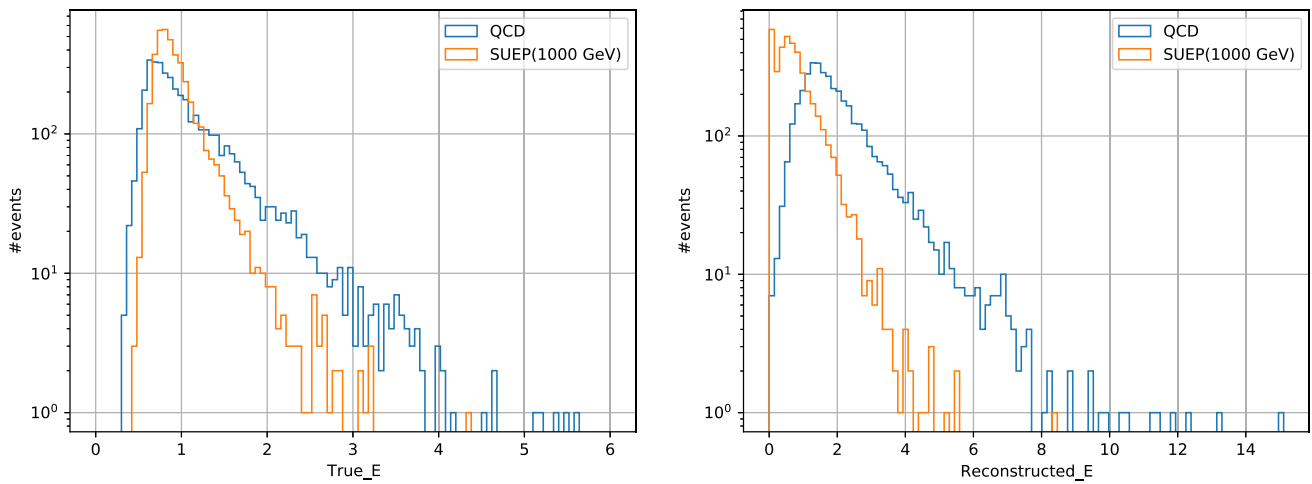
The biggest challenge we faced in accomplishing the task was data sparsity, as mentioned in Sect. 2.2. The model had a huge tendency to learn the zeros rather than the nonzero values. The standard loss functions, such as Mean Squared Error [22] and Cross-Entropy [23], did not work well. We exploited the inverse of the so-called Dice Loss [24] as the loss function, as given by Eq. 1. The  $x_i$  and  $x'_i$  denote the  $i^{\text{th}}$ -pixel values for the input and output, respectively. The  $N$  is the total number of pixels. The denominator term contains the pixel-wise product of input and output, which forces the output to have the same nonzero pixels as the input to minimize the reconstruction loss. On the other hand, the numerator term and minus 1 ensure that the loss function gets its minimum at zero if the output becomes equal to the input.

$$L(x, x') = \frac{\sum_i^N x_i^2 + \sum_i^N x_i'^2}{2 \sum_i^N x_i x_i'} - 1 \tag{1}$$

We used two independent sets of 50k QCD events for the model training and validation. During the model training, no overfitting was observed, and the loss minimization was smooth and saturated after 57 epochs.

**Table 1** The ConvAE architecture

Layer(kernel_size, strides)	Shape	BatchNormalization	Activation function
Input	(288, 360, 3)		
Conv2D((3, 3), (3, 3))	(96, 120, 128)	Yes	PReLU
Conv2D((3, 3), (2, 2))	(48, 60, 64)	Yes	PReLU
Conv2D((3, 3), (2, 2))	(24, 30, 32)	Yes	PReLU
Conv2D((3, 3), (2, 2))	(12, 15, 16)	Yes	PReLU
Conv2D((3, 3), (2, 3))	(6, 5, 8)	Yes	PReLU
Conv2DTranspose((3, 3), (2, 3))	(12, 15, 16)	Yes	PReLU
Conv2DTranspose((3, 3), (2, 2))	(24, 30, 32)	Yes	PReLU
Conv2DTranspose((3, 3), (2, 2))	(48, 60, 64)	Yes	PReLU
Conv2DTranspose((3, 3), (2, 2))	(96, 120, 128)	Yes	PReLU
Conv2DTranspose((3, 3), (3, 3))	(288, 360, 3)	Yes	ReLU

**Fig. 3** Comparisons of true  $E_T$  (left) and reconstructed  $E_T$  (right) (sum over all the pixels) between QCD and SUEP(1000 GeV)

## 4 Results

### 4.1 Autoencoder reconstruction

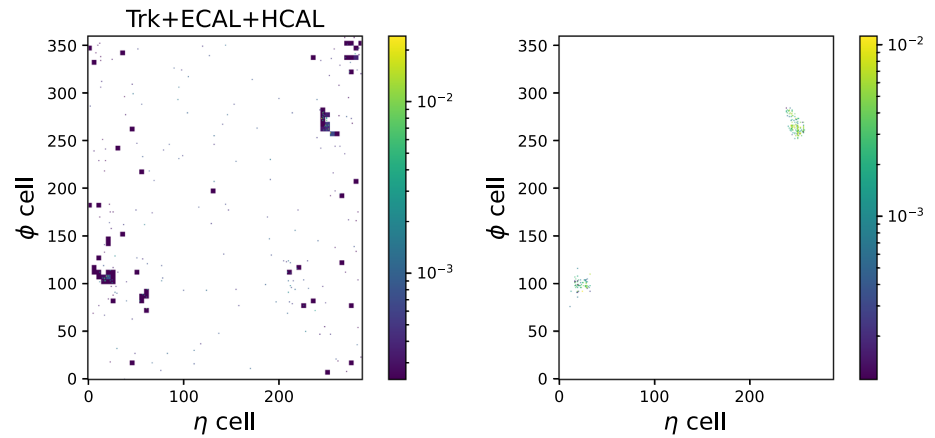
The trained ConvAE is tested using 4k QCD and SUEP events each, which satisfy the event preselection criterion detailed in Sect. 2.2. Figure 3 represents a comparison of true  $E_T$  (left) and reconstructed  $E_T$  (right)—sum over all the pixels—between QCD and SUEP(1000 GeV). It can be noted that the model reconstructs the  $E_T$  for QCD events quite well but struggles for SUEP(1000 GeV)  $E_T$  reconstruction.

The combined true  $E_T$  for a QCD event in the inner tracker, ECAL, and HCAL is given in the left plot of Fig. 4 (same as Fig. 1 but combined now). The reconstructed  $E_T$  for the same event is shown in the right plot of Fig. 4. The model is able to reconstruct the nonzero input pixels quite well. Similar plots for a SUEP(1000 GeV) event are given in Fig. 5. In this case, the model reconstructed only the large  $E_T$  clusters (in the top left quadrant) but failed to reconstruct the whole spherically symmetric signature of a large number of particles in  $\phi$ .

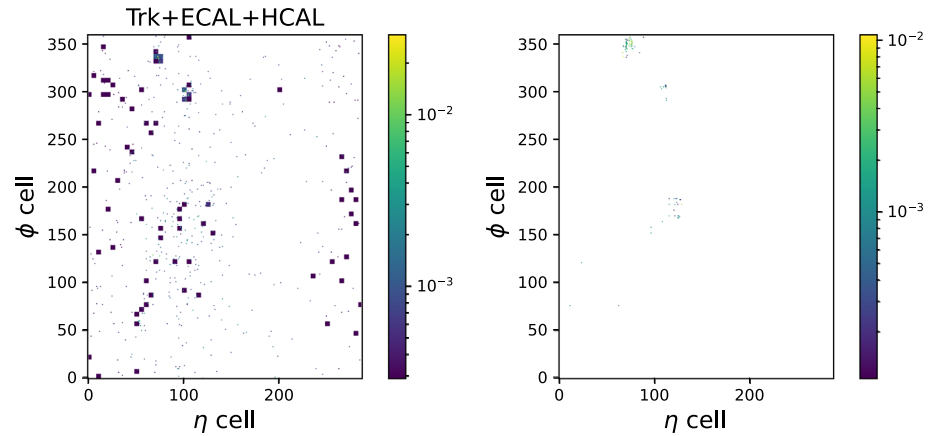
### 4.2 Autoencoder performance

The comparisons of ConvAE reconstruction loss for QCD background and SUEP signal with the considered mediator mass scenarios are given in Fig. 6. The corresponding Receiver Operating Characteristic (ROC) curves are given in Fig. 7. The ConvAE reconstruction loss provides decent discrimination power. The areas under ROC curves (AUCs) are  $\sim 70\%$ , and they barely depend on the mediator

**Fig. 4** Combined true  $E_T$  (left) and reconstructed  $E_T$  (right) for a QCD event in the inner tracker, ECAL, and HCAL



**Fig. 5** Combined true  $E_T$  (left) and reconstructed  $E_T$  (right) for a SUEP(1000 GeV) event in the inner tracker, ECAL, and HCAL



mass. For a target of 40% signal-selection efficiency, the QCD event mistagging rate ranges from 10.3% for SUEP(1000 GeV) to 16% for SUEP(125 GeV).

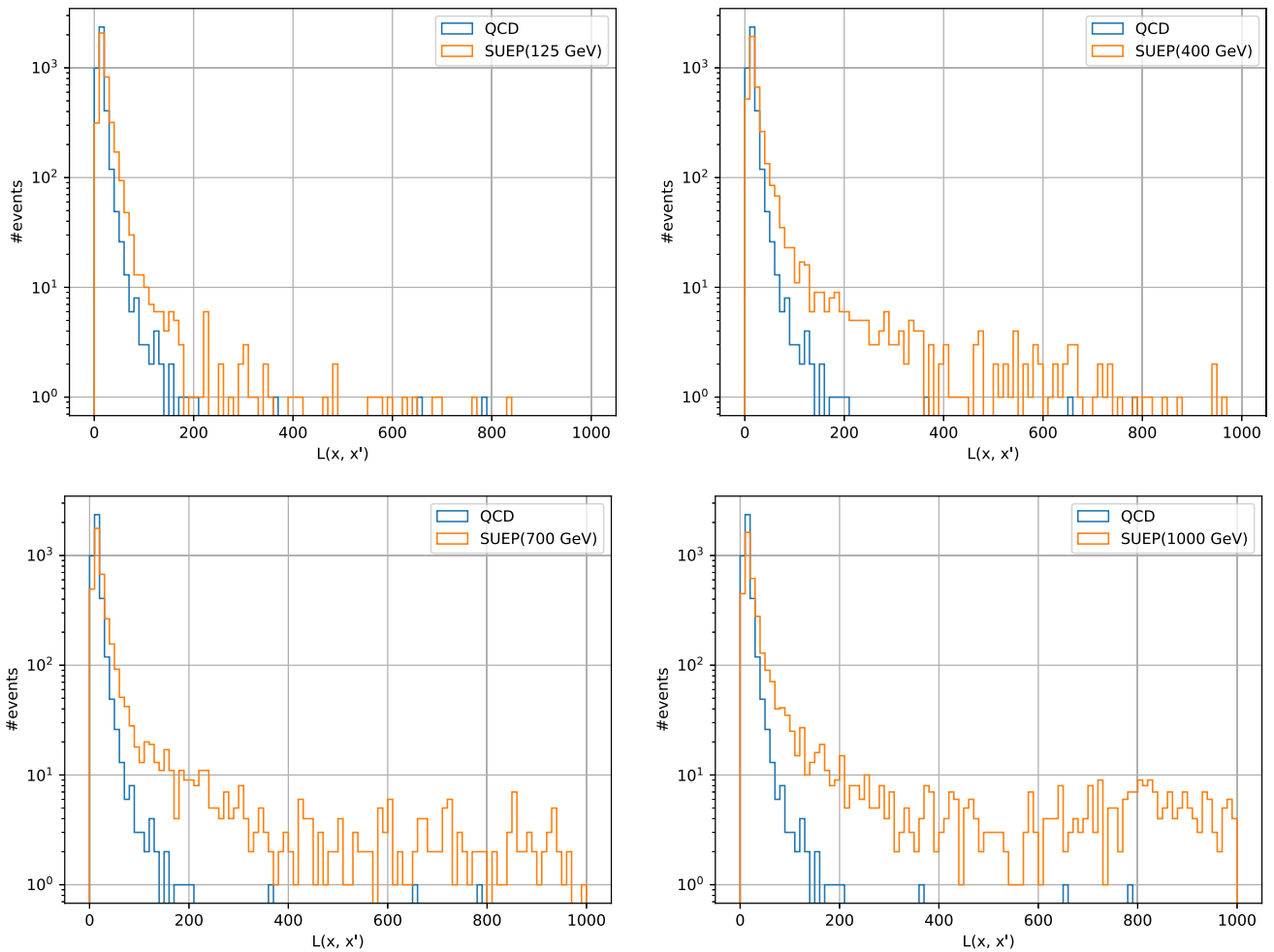
We investigated other ConvAE reconstruction parameters to see if they could be used to further enhance the performance. We observed a parameter, the inverse of reconstructed  $E_T$ , which provides much better discrimination power. The comparisons of the inverse of the reconstructed  $E_T$  for QCD background and SUEP signal with the considered mediator mass scenarios are given in Fig. 8, and the corresponding ROC curves are given in Fig. 9. The AUCs range from 75.6% for SUEP(125 GeV) to 86.9% for SUEP(1000 GeV). For a target of 40% signal-selection efficiency, the QCD mistagging rate ranges from 2% for SUEP(1000 GeV) to 11.9% for SUEP(125 GeV).

The model inference time has been quantified using a CPU device, an Intel® Core™ i5-9600KF processor [25], and found to be  $\sim 20$  ms, which is very well within the processing-time limits of the HLT system ( $\mathcal{O}(10^2)$  ms).

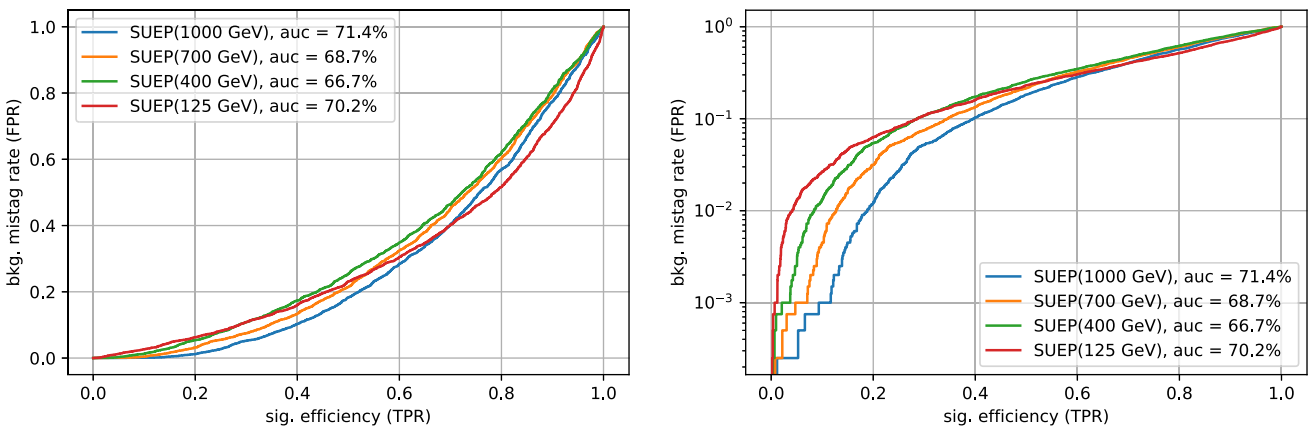
## 5 Conclusions

A deep convolutional neural autoencoder network, trained using background events by taking their signatures in different sub-detectors of an experiment at the Large Hadron Collider as  $n$ -channel image data, is highly effective as an Anomaly Detection (AD) technique. The training of the autoencoder using unlabeled background events makes this technique highly model-agnostic. We have developed an autoencoder for AD with the Compact Muon Solenoid experiment, which can easily be adapted for other Large Hadron Collider experiments. We have exploited the raw reconstructed signature in the High-Level Trigger system, with the clear benefit of not relying on online particle reconstruction, which could fail to reconstruct new physics signatures.

The autoencoder has been trained using QCD events by taking transverse energy deposits in the inner tracker, electromagnetic calorimeter, and hadron calorimeter sub-detectors as 3-channel image data. The biggest challenge was the sparsity of the data: only  $\sim 0.5\%$  of the total  $\sim 300k$  pixels per image had nonzero values. The model had a high tendency to learn zeros rather than nonzero values. We explored several loss functions, and ultimately, a nonstandard loss function served the purpose—the inverse of the so-called Dice Loss, which contains the pixel-wise multiplication of input and output that forces the output images to have the same nonzero pixels as the input.

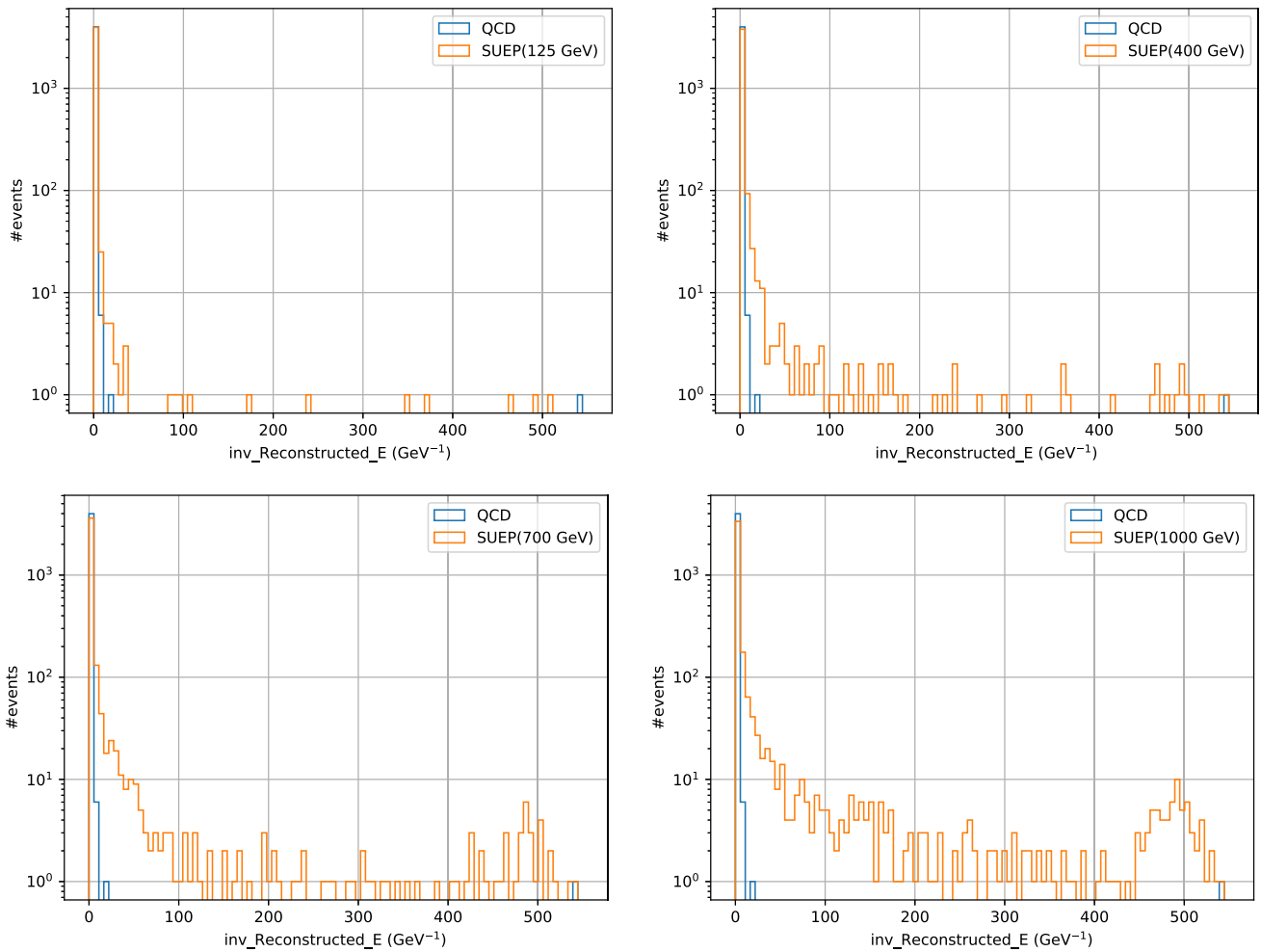


**Fig. 6** Comparisons of ConvAE reconstruction loss for QCD and SUEP(125 GeV) (top left), SUEP(400 GeV) (top right), SUEP(700 GeV) (bottom left), and SUEP(1000 GeV) (bottom right)

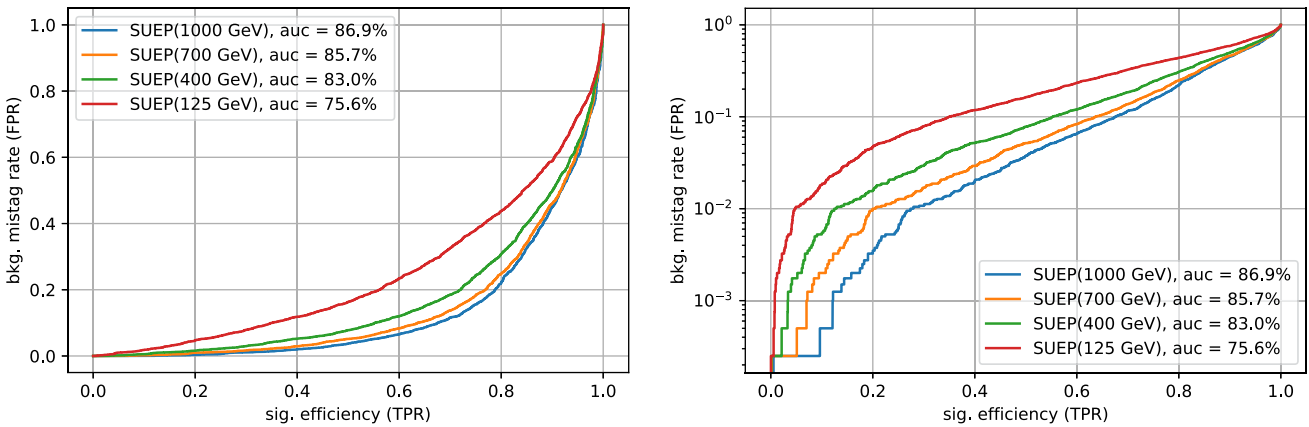


**Fig. 7** ROC curves for ConvAE reconstruction loss as an AD proxy. The right plot is the same as the left plot but with a logarithmic y-scale

The trained model has been tested for Soft Unclustered Energy Patterns (or SUEPs) detection—a new physics signal with an experimental signature of high-multiplicity spherically symmetric Standard Model particles, anomalous to QCD jets. To check the robustness of the autoencoder, we considered several signal scenarios depending on the new physics model. It has been observed that the model can detect 40% of the SUEP events at the expense of a QCD event mistagging rate as low as 2%. The model inference time using the Intel® Core™ i5-9600KF processor is measured to be  $\sim 20$ ms, which perfectly satisfies the HLT system’s latency requirements.



**Fig. 8** Comparisons of the inverse of reconstructed  $E_T$  for QCD and SUEP(125 GeV) (top left), SUEP(400 GeV) (top right), SUEP(700 GeV) (bottom left), and SUEP(1000 GeV) (bottom right)



**Fig. 9** ROC curves for the inverse of reconstructed  $E_T$  as an AD proxy. The right plot is the same as the left plot but with a logarithmic y-scale

**Acknowledgements** S. S. Chhibra, N. Chernyavskaya, and B. Maier are supported by CERN, Switzerland under the Senior Research Fellowship Program. S. S. Chhibra is partly supported by QMUL, UK. B. Maier is partly supported by KIT, Germany. M. Pierini is supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (Grant Agreement No. 772369). S. Hasan is supported by SNS di Pisa, Italy and partly supported by ETH Zurich, Switzerland.

**Funding** HORIZON EUROPE Framework Program (772369).

**Data Availability** The data used in the creation of this manuscript will be made available upon request.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. M.J. Strassler, K.M. Zurek, Echoes of a hidden valley at hadron colliders. *Phys. Lett. B* **651**, 374–379 (2007)
2. S. Knapen, S.P. Griso, M. Papucci, D.J. Robinson, Triggering Soft Bombs at the LHC. *J. High Energy Phys.* **08**, 076 (2017)
3. J. Barron, D. Curtin, G. Kasieczka, T. Plehn, A. Spourdalakis, Unsupervised hadronic SUEP at the LHC. *J. High Energy Phys.* **12**, 129 (2021)
4. <https://www.darkmachines.org/>
5. <https://lhco2020.github.io/homepage/>
6. V. Khachatryan et al., The CMS trigger system. *JINST* **12**(01), P01020 (2017)
7. S. Chatrchyan et al., The CMS experiment at the CERN LHC. *JINST* **3**, S08004 (2008)
8. D. Kar, S. Sinha, Exploring jet substructure in semi-visible jets. *SciPost Phys.* **10**(4), 084 (2021)
9. P. Schwaller, D. Stolarski, A. Weiler, Emerging jets. *J. High Energy Phys.* **05**, 059 (2015)
10. T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C.O. Rasmussen, P.Z. Skands, An introduction to PYTHIA 8.2. *Comput. Phys. Commun.* **191**, 159–177 (2015)
11. J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, M. Selvaggi, DELPHES 3, a modular framework for fast simulation of a generic collider experiment. *J. High Energy Phys.* **02**, 057 (2014)
12. A. Tumasyan et al., Search for resonant production of strongly coupled dark matter in proton-proton collisions at 13 TeV. *J. High Energy Phys.* **06**, 156 (2022)
13. S. Knapen, J. Shelton, X. Dong, Perturbative benchmark models for a dark shower search program. *Phys. Rev. D* **103**(11), 115013 (2021)
14. I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, Cambridge, 2016)
15. K. O'Shea, R. Nash, An introduction to convolutional neural networks. *CoRR*, abs/1511.08458 (2015)
16. V. Dumoulin, F. Visin, A guide to convolution arithmetic for deep learning (2016). <http://arxiv.org/abs/1603.07285> arxiv:1603.07285
17. K. He, X. Zhang, S. Ren, S. Jian, Delving deep into rectifiers: surpassing human-level performance on imagenet classification. *Proceedings of the IEEE International Conference on Computer Vision* (2015), p. 1026–1034
18. F. Chollet et al., Keras (2015)
19. M. Abadi et al., TensorFlow: large-scale machine learning on heterogeneous systems, software available from tensorflow.org (2015)
20. D.P. Kingma, J.B. Adam, A method for stochastic optimization (2014). Cite [arxiv:1412.6980](https://arxiv.org/abs/1412.6980). Comment: Published as a Conference Paper at the 3rd International Conference for Learning Representations, San Diego (2015)
21. <https://www.nvidia.com/en-us/data-center/v100/>
22. C. Sammut, G.I. Webb (eds.), *Encyclopedia of Machine Learning and Data Mining*, 2nd edn. (Springer, New York, 2017)
23. P. De Boer, K.S. Mannor, R.Y. Rubinstein, A tutorial on the cross-entropy method. *Ann. Oper. Res.* **134**, 19–67 (2004)
24. R. Deng, C. Shen, S. Liu, H. Wang, X. Liu, Learning to predict crisp boundaries, in *Computer Vision—ECCV 2018*. ed. by V. Ferrari, M. Hebert, C. Sminchisescu, Y. Weiss (Springer, Cham, 2018), pp.570–586
25. <https://www.intel.com/content/www/us/en/products/sku/190884/intel-core-i59600kf-processor-9m-cache-up-to-4-60-ghz/specifications.html>