Masked Image Modeling for Image Completion on Simulated Calorimeter Data

Kristina Jaruskova[1,2], R. Cardoso[2], S. Vallecorsa[2], P. Raikwar[2], A. Zaborowska[2]

[1]Czech Technical University in Prague, [2]CERN

jaruskri@fjfi.cvut.cz

Introduction

- Monte Carlo simulations of calorimeter showers are very time demanding and finding a faster and versatile alternative is of a high interest.
- Masked modeling is an established technique for unsupervised representation learning achieving remarkable results in the NLP domain.
- Can the masked modeling be used to learn rich representations of calorimeter showers?
- We present an adaptation of the masked modeling with transformer network to calorimeter simulations. We trained our model on general calorimeter showers recorded in a cylindrical readout mesh.

Training Data

- ECAL showers induced by a single electron (available on `Zenodo.org` [1]).
- Initial particle energies: 64 GeV, 128 GeV, and 256 GeV.
- Energy depositions recorded in a cylindrical readout mesh with $15 \times 50 \times 24$ cells in the $R \times \varphi \times Z$ coordinates.
- High dynamic range of pixel values spanning over the $10^{-4} - 10^{2}$ GeV range.

Preprocessing Images

- Scaling cell energies:
  - The cell energies are divided by the initial particle energy.
  - All values are scaled to the range [0, 1].
- Converting an image to a sequence:
  - Following the ViT example [2].
  - Dividing the cylindrical mesh into smaller segments (*patches*).
  - Splitting in R and Z direction only because of the rotational symmetry of the shower images.
- Embedding individual patches:
  - Embedding each patch into lower-dimensional space, as well as the patch positions.
  - Adding the patch and position embedding vectors (element-wise sum).

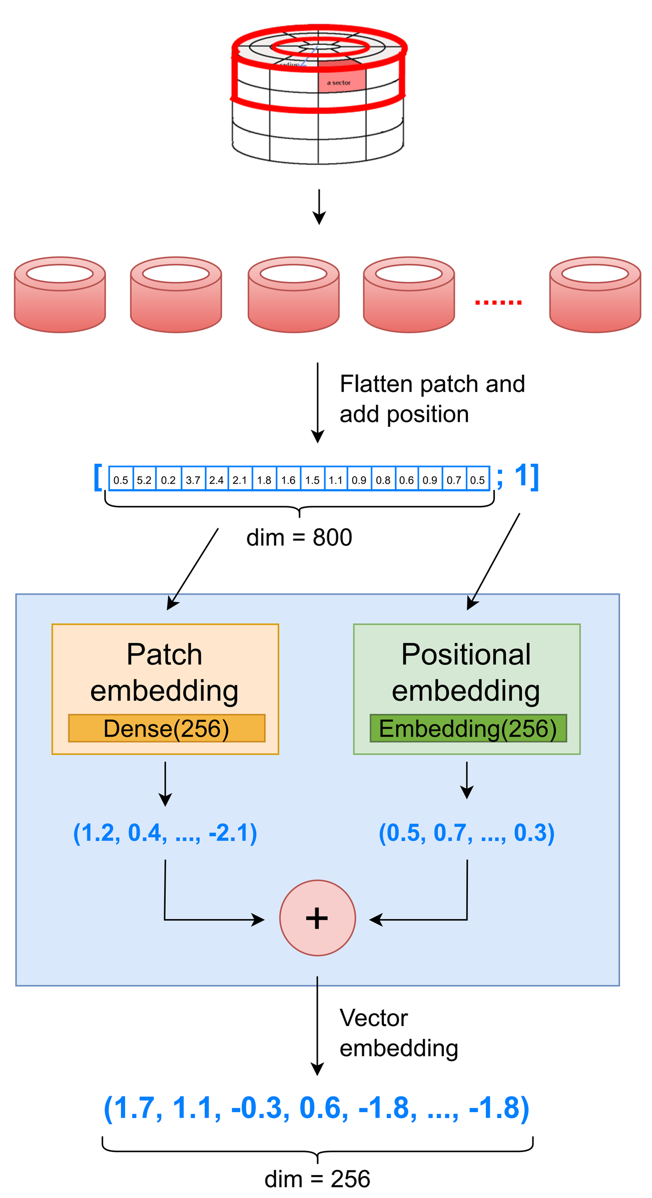


Fig. 1: Image-to-sequence transformation.

Model Structure

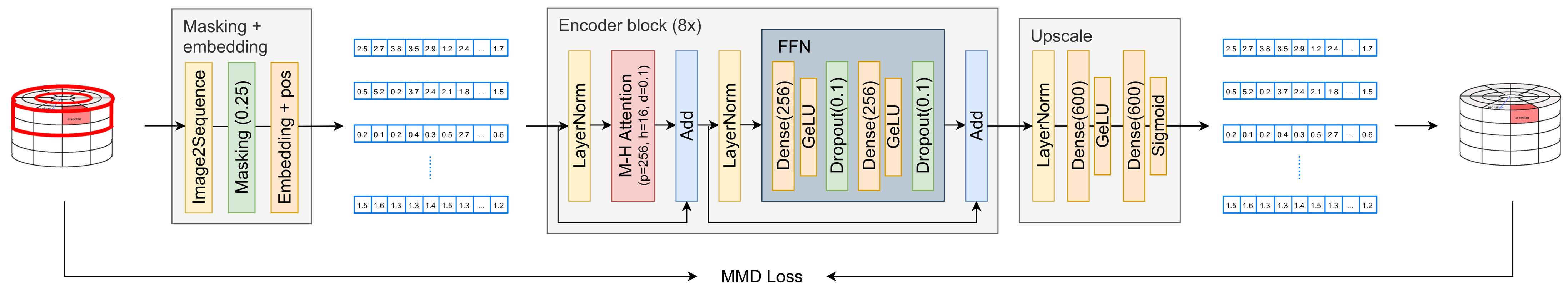


Fig. 2: Complete design of the model.

Image preprocessing First, the shower images are divided into patches to create a sequence. Afterwards, random patches are masked (all values replaced by zeros) and all patches are transformed into the embedding space.

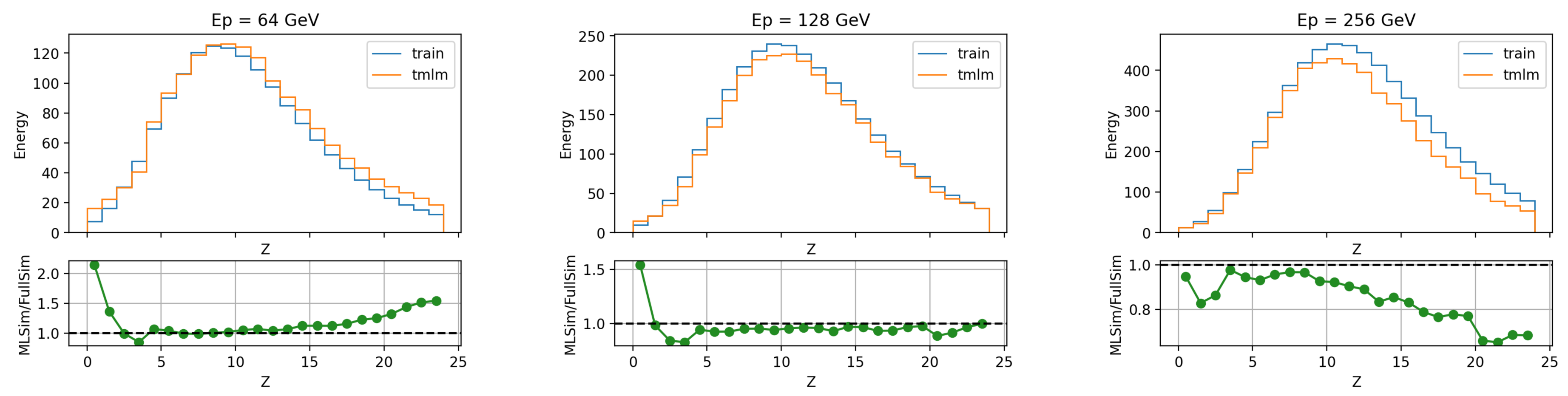
Reconstruction The encoder-like network with self-attention blocks uses bi-directional information from the input sequence to reconstruct masked patches. New patch representations are upsampled to the original dimension and the cylindrical mesh is reassembled.

Training The embedding layers loop over patches, while the encoder loops over images (sequences of embedded patches). For this reason, each batch is split into mini-batches of 8 images and the weights in the first part of the model are adjusted after each mini-batch.

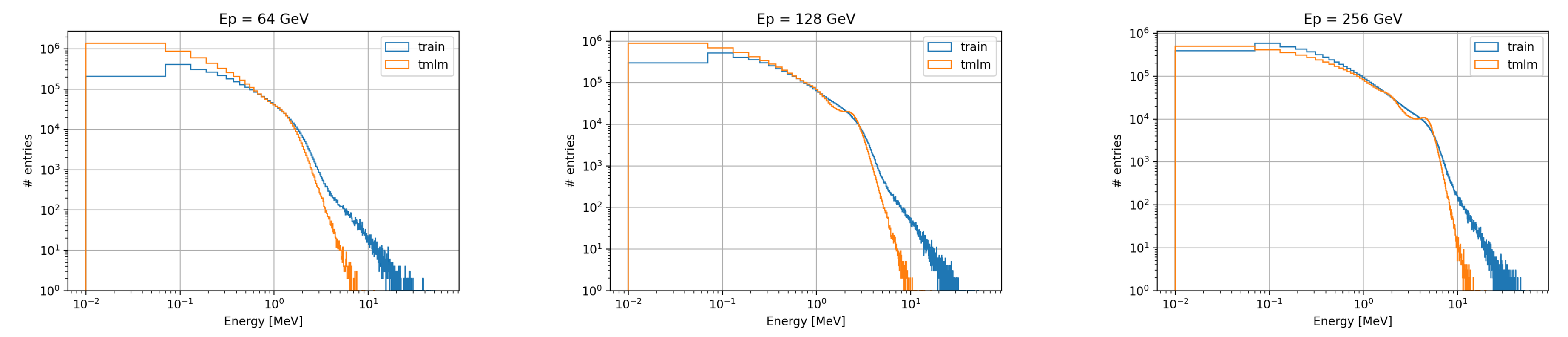
Loss The quality of reconstructed images is measured by the MMD loss (maximum mean discrepancy) with the multi-scale kernel.

Results

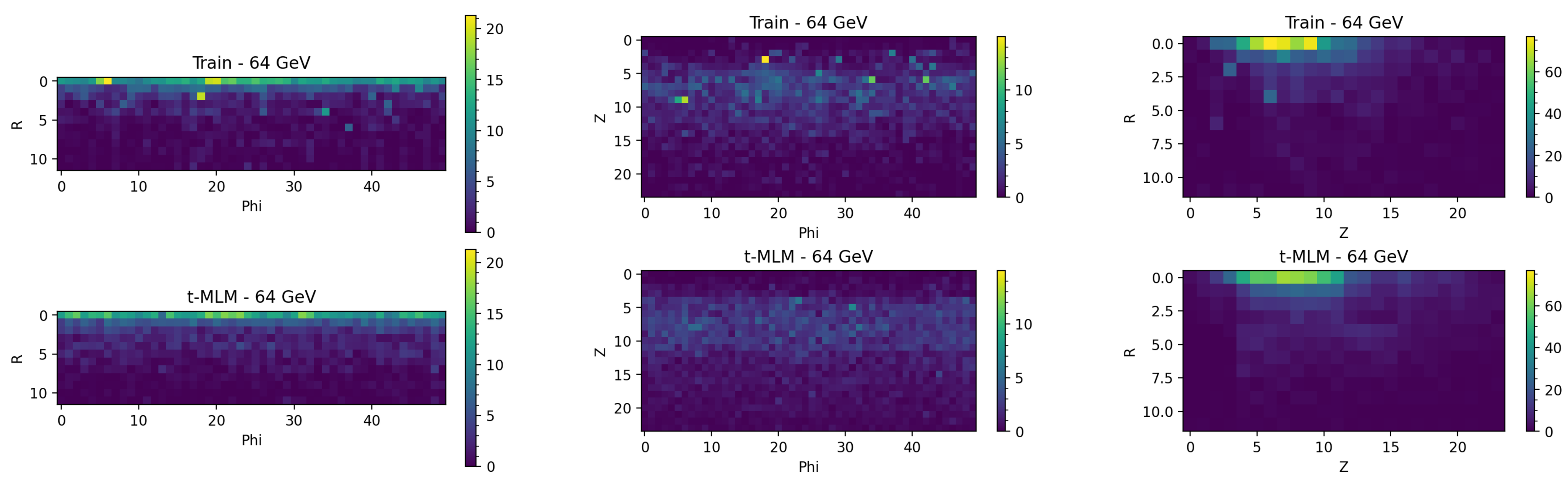
Shower shapes Depiction of average profiles along the main axis Z of the cylinder. The model reconstructs the showers shapes well for all tested initial particle energies without explicitly providing its value.



Cell energies Histograms of the cell energies, visualised in a $\log_{10}$ scale. We observe solid results for $10^{-1} - 10^{1}$ GeV cells. The model currently struggles to replicate the high-energy depositions that are rare.



2D plots Example of 2-dimensional projections of energy depositions for one shower induced by a 64 GeV electron. Overall, energy distributions are correctly replicated and the plots have the desired granularity. We observe a surplus of low-energy cells instead of empty cells.



Conclusion

- We demonstrated that a transformer-based model can learn good representations of ECAL showers using the masked modeling approach.
- The model is able to adapt to different values of primary particle energy based solely on the input embeddings and generate showers with granularity in individual cells.

References

[1] Zaborowska A., Salamani D. *High Granularity Electromagnetic Calorimeter Shower Images*, 2022. `zenodo.org/records/6082201`

[2] Dosovitskiy A., et al. *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*, 2020. `https://arxiv.org/abs/2010.11929`