

Evaluation and Optimisation of a Generative-Classification Hybrid Variational Autoencoder in the Search for Resonances at the LHC

Finn Stevenson¹, Benjamin Lieberman¹, Abhaya Swain¹, Bruce Mellado^{1,2}

¹School of Physics and Institute for Collider Particle Physics, University of the Witwatersrand, Johannesburg, Wits 2050, South Africa

²iThemba LABS, National Research Foundation, PO Box 722, Somerset West 7129, South Africa

E-mail: finn.david.stevenson@cern.ch

Abstract. The Standard Model (SM) of particle physics was completed by the discovery of the Higgs boson in 2012 by the ATLAS and CMS collaborations. However, the SM is not able to explain a number of phenomena and anomalies in the data. These discrepancies to the SM motivate the search for new bosons. In this paper, searches for new bosons are completed by looking for $Z\gamma$ resonances in $Z\gamma$ ($pp \rightarrow H \rightarrow Z\gamma$) fast simulation events. This research makes use of a Variational Autoencoder (VAE), in the search for new bosons. The functionality of a VAE to be trained as both a generative model and a classification model makes the architecture an attractive option for aiding the search. The VAE is used as a generative model to increase the amount of $Z\gamma$ fast simulation Monte Carlo data whilst simultaneously being used to classify samples containing injected signals that differ from the Monte Carlo data on which the model was trained. This work concentrates on the final evaluation and optimisation of the VAE for the generative task.

1 Introduction

The Standard Model of particle physics was completed by the discovery of the Higgs boson in 2012 by the ATLAS and CMS collaborations. [1, 2] However the SM is not able to explain a number of phenomena and anomalies in the data. These discrepancies to the SM motivate the search for new bosons [3, 4, 5, 6]. In this greater study, searches for new bosons are completed by looking for Z resonances in $Z\gamma$ ($pp \rightarrow H \rightarrow Z\gamma$) fast simulation events. This research makes use of a Variational Auto encoder (VAE) to aid in the search for new bosons. In high energy physics resonance searches where semi-supervised machine learning models are used, a frequentist approach can be utilised to evaluate the extent of background events falsely labeled as signal events. This approach requires large amounts of simulated data. Deep Learning models such as a VAE can be used in frequentist studies for data generation. The functionality of a VAE to be trained as both a generative model and a classification model makes the architecture an attractive option for aiding the search. The VAE is used as a generative model to increase the amount of $Z\gamma$ fast simulation Monte Carlo data whilst simultaneously being able to classify samples containing injected signals that differ from the Monte Carlo data on which the model was trained.

2 Methodology

A VAE is a encoder-decoder type neural network used for both data generation and classification tasks. Compared to the standard Auto-encoder (AE), architectural changes and an additional component added to the loss function regularise training and improve the generative capability of the model by ensuring appropriate latent space properties. A VAE is trained to minimise the loss between the input data (kinematic variable event) and the encoded-decoded output (reconstructed event). The input is encoded as a distribution over the latent space of the VAE (latent space variables are forced to be Gaussian). This allows for some regularisation of the latent space.

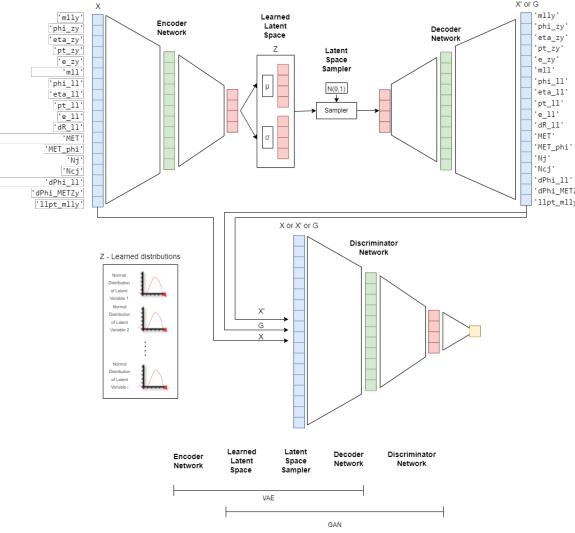


Figure 1. Diagram of VAE Base and greater VAE+D model Architecture, showing encoder network, decoder network, learned latent space and discriminator network.

The loss function that is minimised when training a VAE is composed of a reconstruction loss component that is responsible for forcing the output of the decoder to be as close to the input, and secondly, a regularisation loss component, that serves to regularise the organisation of the latent space by making the distributions returned by the encoder close to a standard normal distribution.

$$L_{VAE} = L_R + \beta_V * L_{KL} \quad (1)$$

$$L_R = \overline{((X' - X)^2)} \quad (2)$$

$$L_{KL} = \sum (\sigma^2 + \mu^2 - \log \sigma - 1) \quad (3)$$

Where X is the input event and X' is the reconstructed event. The addition of a discriminator and the notion of adversarial training helps in the training of the VAE encoder-decoder network. Figure 1 shows the architecture diagram of the VAE+D model. The VAE+D loss functions seen below are slightly different to the VAE, whilst still including the main loss components of the original VAE. Unlike the VAE, the VAE+D has a loss function for each individual network, the encoder, decoder and discriminator and each network's weights are updated individually at different times during a forward pass of the overall VAE+D. Similar to a Generative Adversarial Network (GAN), the discriminator and the VAE are trained simultaneously, with the discriminator learning to distinguish fake events from real events and the VAE learning to reproduce real events accurately. The VAE+D loss functions and components are as follows:

$$L_{disc} = BCE_{real} + BCE_{recon} + BCE_{gen} \quad (4)$$

$$L_{dec} = \gamma * Loss_R - Loss_{disc} \quad (5)$$

$$L_{enc} = Loss_{KL} + \gamma * Loss_R \quad (6)$$

Where BCE is the binary cross entropy between either the actual data against actual data, reconstructed data or generated data. The final discriminator loss function, L_{disc} is obtained as the sum of the three BCE based losses. γ is a coefficient of the reconstruction loss term in the loss function. VAEs have many hyper-parameters that can be optimised in order to achieve the best model. This hyper-parameter optimisation can be done using a variety of methodologies and available libraries, however in this work a manual optimisation loop was created.

3 Results

The Results of the addition of the discriminator network to the VAE can be seen in the figures below. It can be seen that the optimised VAE+D model is a better generative model for the chosen $Z\gamma$ data.

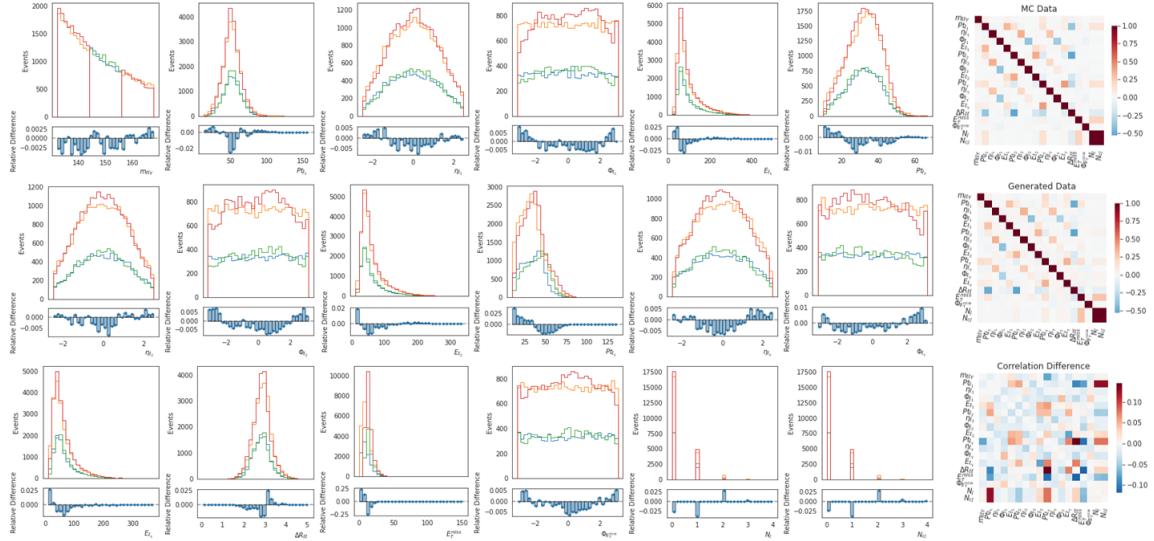


Figure 2. Distribution and Correlation Plots of Generated Event Features vs. MC Data for VAE+D Model

References

- [1] Aad G *et al.* (ATLAS) 2012 *Phys. Lett. B* **716** 1–29 (*Preprint* 1207.7214)
- [2] Chatrchyan S *et al.* (CMS) 2012 *Phys. Lett. B* **716** 30–61 (*Preprint* 1207.7235)
- [3] von Buddenbrock S, Cornell A S, Fadol A, Kumar M, Mellado B and Ruan X 2018 *J. Phys. G* **45** 115003 (*Preprint* 1711.07874)
- [4] Buddenbrock S, Cornell A S, Fang Y, Fadol Mohammed A, Kumar M, Mellado B and Tomiwa K G 2019 *JHEP* **10** 157 (*Preprint* 1901.05300)
- [5] von Buddenbrock S, Ruiz R and Mellado B 2020 *Phys. Lett. B* **811** 135964 (*Preprint* 2009.00032)
- [6] Hernandez Y, Kumar M, Cornell A S, Dahbi S E, Fang Y, Lieberman B, Mellado B, Monnakgotla K, Ruan X and Xin S 2021 *Eur. Phys. J. C* **81** 365 (*Preprint* 1912.00699)