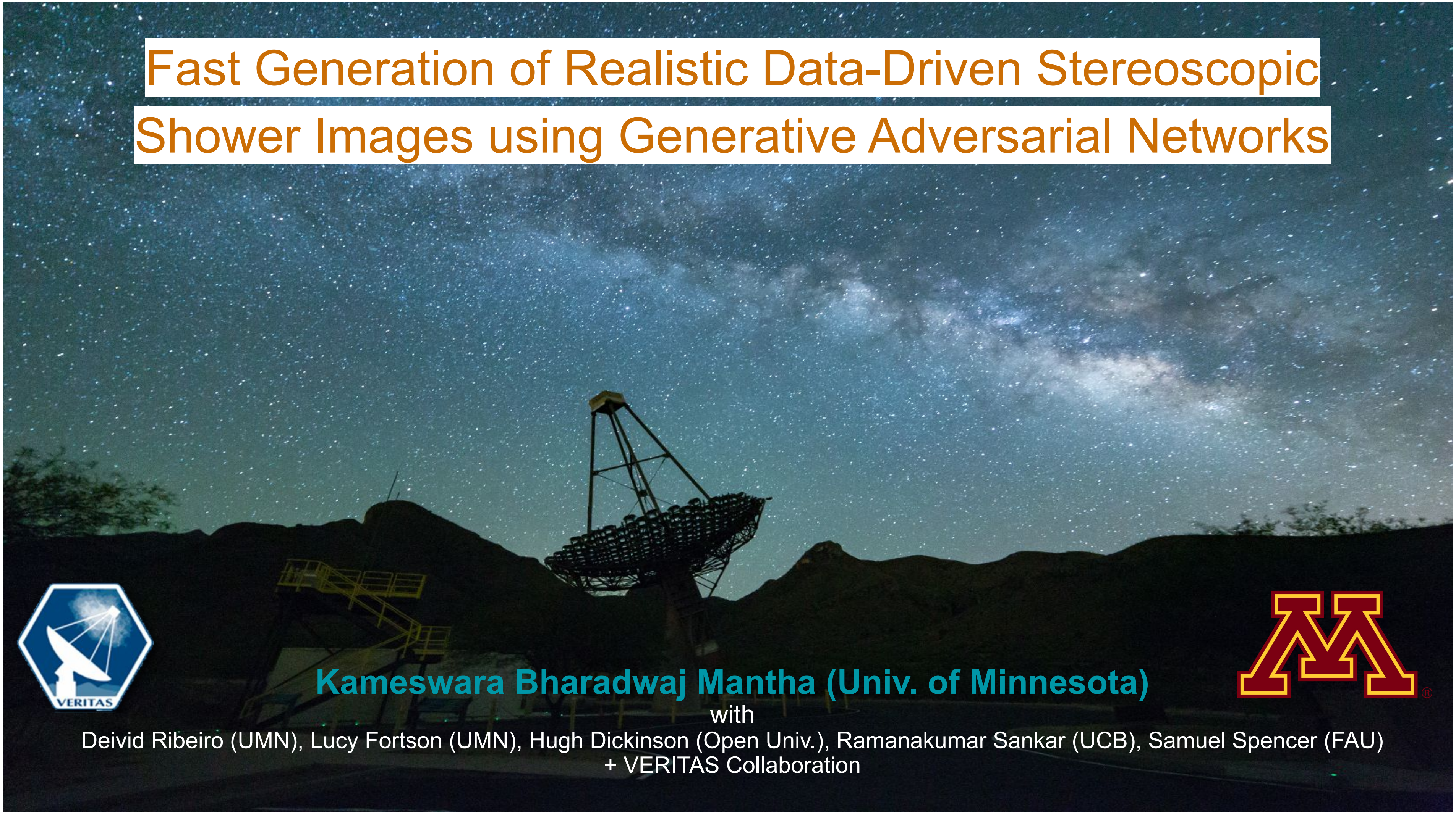
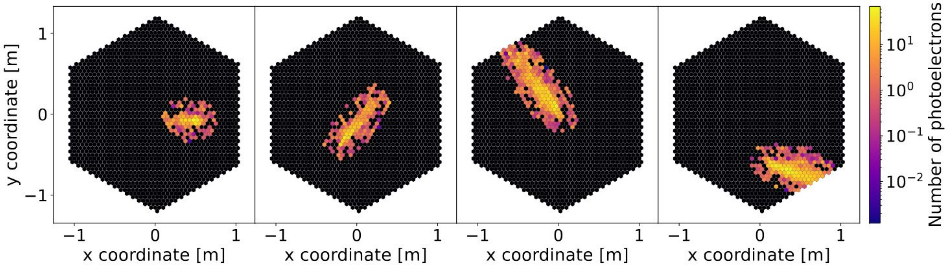


Fast Generation of Realistic Data-Driven Stereoscopic Shower Images using Generative Adversarial Networks

Kameswara Bharadwaj Mantha (Univ. of Minnesota)
with
Deivid Ribeiro (UMN), Lucy Fortson (UMN), Hugh Dickinson (Open Univ.), Ramanakumar Sankar (UCB), Samuel Spencer (FAU)
+ VERITAS Collaboration

Motivation

- **Overarching Motivation:** Tackling **Gamma/Hadron separation** problem in an observational data-driven manner (previous successes have been based on simulated events)
  - Previous efforts have used **supervised ML methods** trained on simulations (*Shilon+18, Miener+22, Spencer+21 Thesis*);
  - Generalized Simulation → Data application can be challenging (domain adaptation can be done)
  - Gamma : Hadron ratio in observational data is **heavily imbalanced (1:1000 - 1:10000)**.
    - Gammas can be perceived as “unusual” (or anomalous) in a pool of hadrons.
- **Unsupervised Deep Learning:** A promising (alternative) approach to learn features in a data-driven way. They are also useful for doing **anomaly detection**!
  - Autoencoders – focused on image-to-image reconstructions
  - **Generative Adversarial Networks** – focused on generation of images that are representative; (e.g., *Elflein+23*)
  - Transformers (and more…)



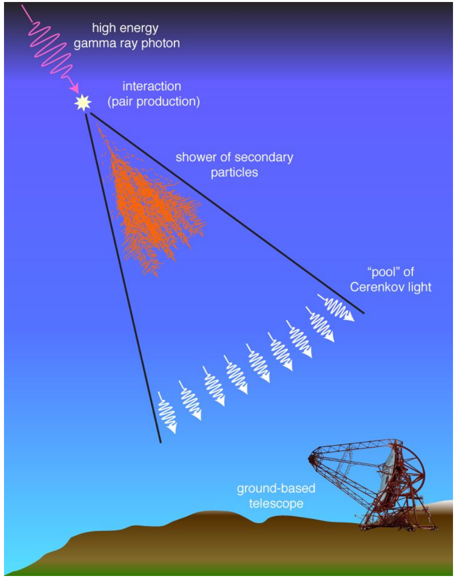
Generated images using GAN from Elflein+23

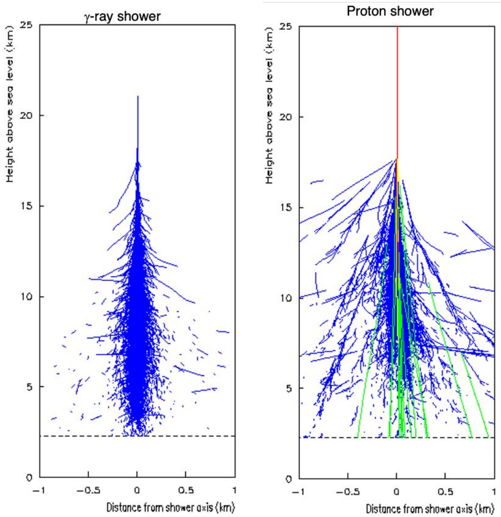
Also see slides from Glombitza and Vuillaume talks (Tue)

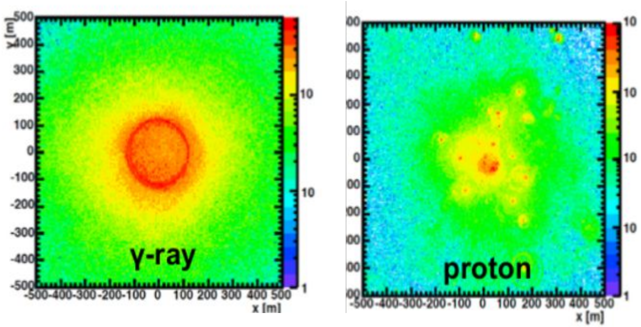
Outline

- **Generative Adversarial Networks (GANs)** are established method to learn image-level features in a data-driven way (*Goodfellow+14*):
  - However - Standard GAN framework poses following issues
    - Training instability: This motivates a new variant Wasserstein GAN (wGAN).
    - Not grounded by physical parameters: To be useful for domain applications, requires incorporation of relevant physical parameters, motivating Conditional wGAN (CwGAN).
    - Previous efforts using CwGANs have been with simulated data.
  - **Our goal**: Train in a data driven manner a CwGAN that ultimately lead us to effective gamma/hadron separation
- **An Intermediate (useful) by-product of CwGAN**: Fast Generation of Realistic Data-Driven Stereoscopic Shower Images - this is the focus of my talk
- **Conclusions** of our CwGAN experiments!

Gamma-Hadron Separation: A Critical Scientific Analysis Step



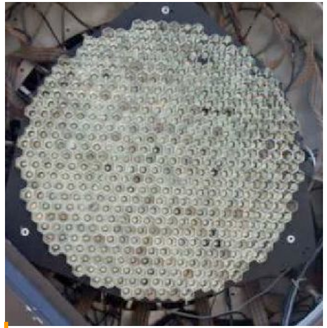




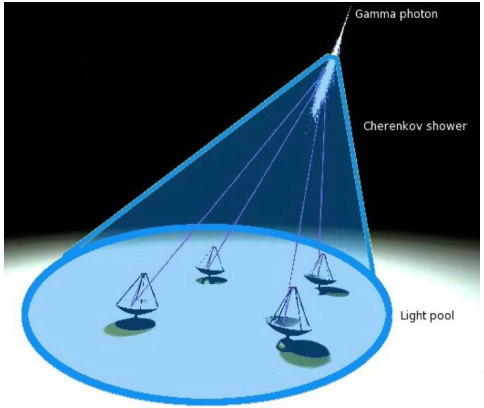
1. Atmospheric showers from gamma rays and cosmic rays create distinct patterns on imaging.
2. We use data from the VERITAS observatory

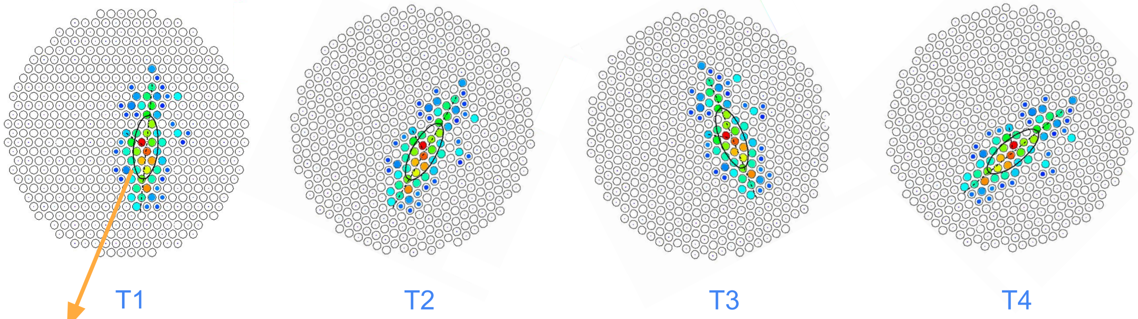
Schematic Overview of the VERITAS Data & Derived Parameters

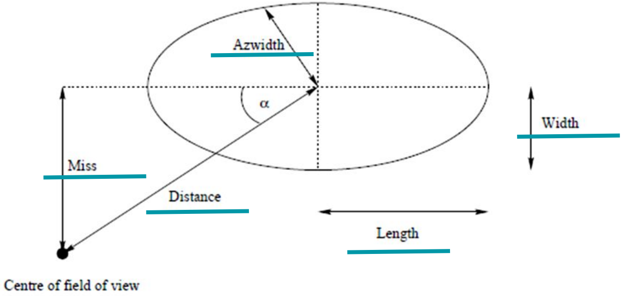
Camera:
499 PMTs
3.5 deg FOV
Hexagonal Pixels



We specifically use stereoscopic, integrated images of all 4 telescopes & their corresponding Hillas Parameters



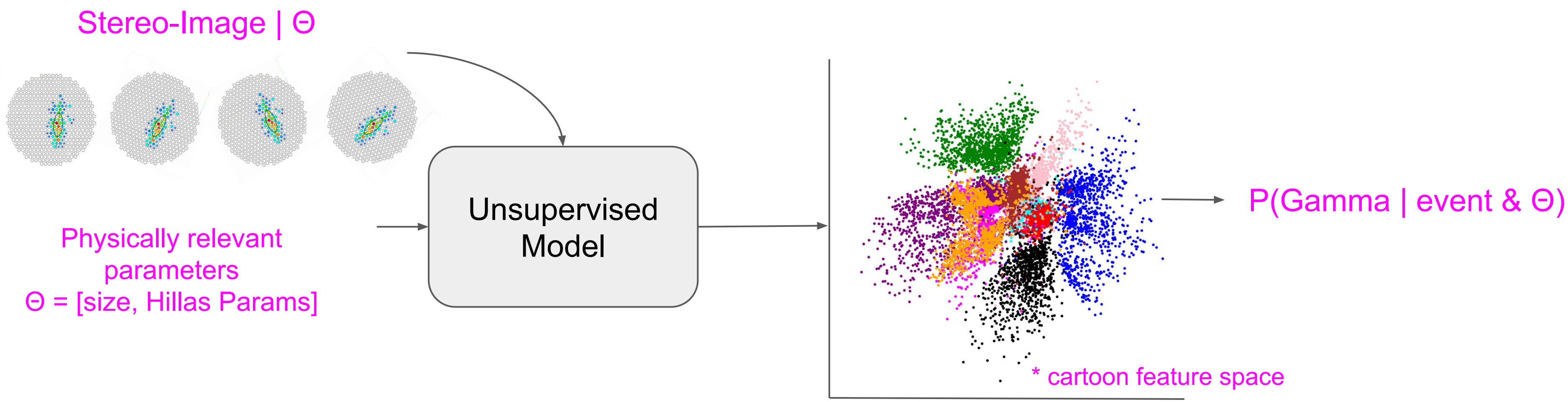




Hillas Parameters are effectively image moments quantifying the captured shower

Schematic of our overarching goal: Data-driven learning of shower-image features

1. We expect that an unsupervised approach helps learn generalized diversity of shower features.
2. We anticipate the data to govern the model on what features are representative.

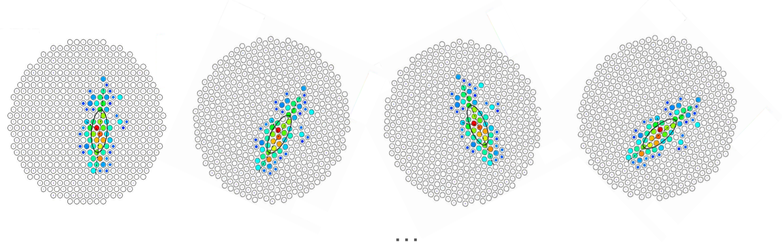


Our goal is to create a model – trained on stereoscopic images alongside physically-relevant parameters – that could learn a lower-dimensional representation which can help better distinguish Gamma vs. Hadrons.

We use VERITAS observational data of the Crab Nebula for our Deep Learning training

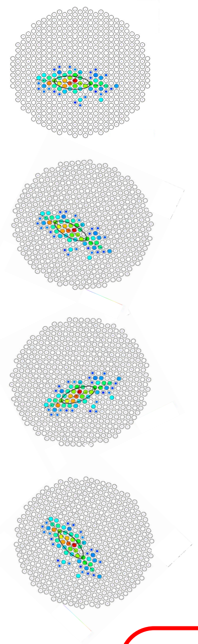
- Crab Nebula (SN 1054)
- ~6500 Lightyears away
- Bright and steady source in TeV.
- **Calibration standard for IACTs.**

Data Used for our Proof-of-Concept Training



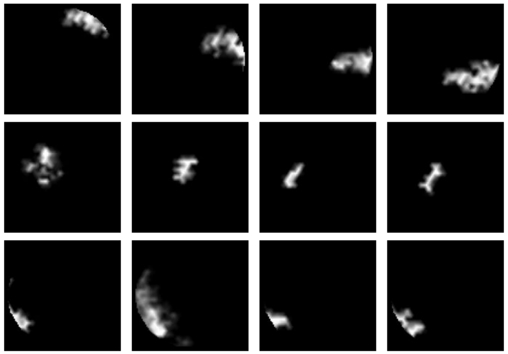
- 1 Crab Run
- ~120K events
- Size (per telescope) > 100

How we make VERITAS data compatible with deep learning tools



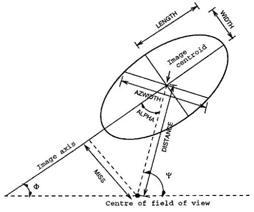
- Hex to rectangular pixel interpolation
- Noise cleaning
- Normalization per Tel [0,1]

We used cleaned image(s) +
Hillas parameters in our work



T1 T2 T3 T4

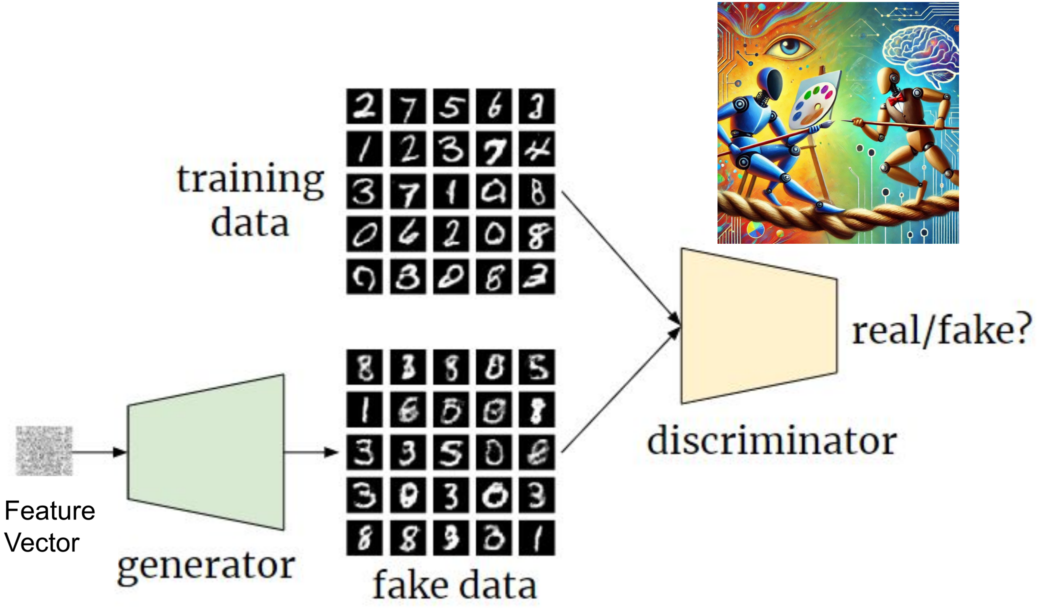
Hillas Parameters
per Telescope



1. We project native hex pixels onto a 96x96 rectangular grid using bi-cubic interpolation.
2. We normalize per-telescope image such that the native charge distribution is conserved.
3. **Normalizing scheme makes a BIG DIFFERENCE in training outcomes!!**

Generative Adversarial Networks (GANs): A good architecture for data-driven feature learning

- An established **unsupervised** feature learning framework
- A GAN, under the hood is comprised of **two CNN-based frameworks** – Generator & Discriminator
- It learns **generalized feature representations** described by the data.
- One can use the representations to **generate realistic synthetic data** and finding **rarely occurring samples**

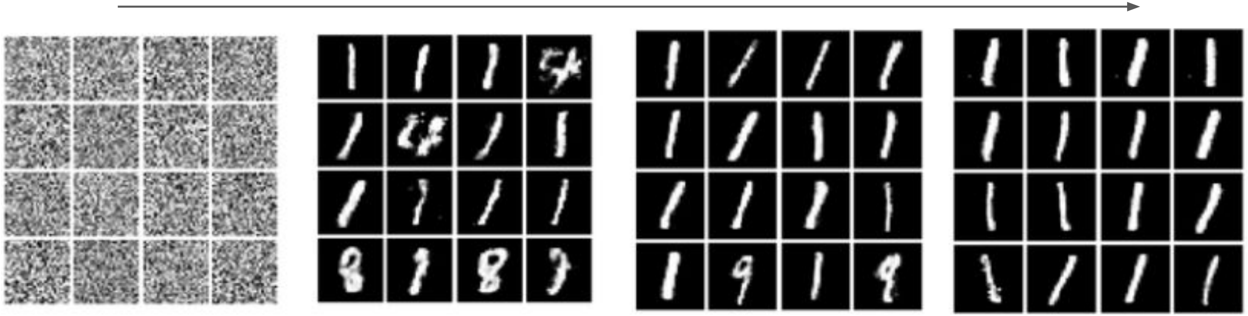


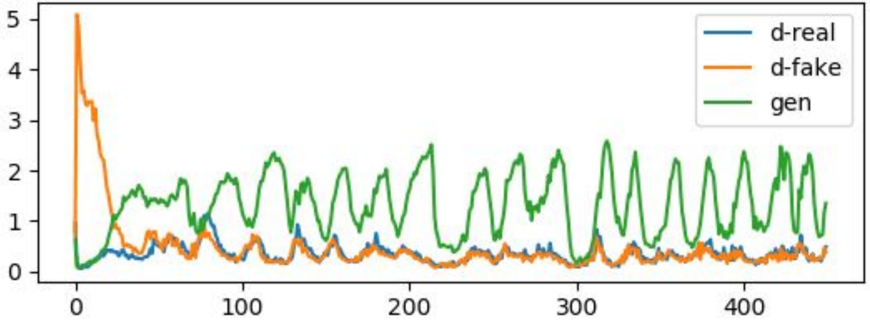
Example synthetically Generated Images

Traditional GANs suffer from training instabilities

1. The Generator can cheat the Discriminator by generating one particularly “good” sample (mode collapse)
2. The learning landscape can become “too volatile” resulting in the Generator to learn-unlearn iteratively.

Training progress



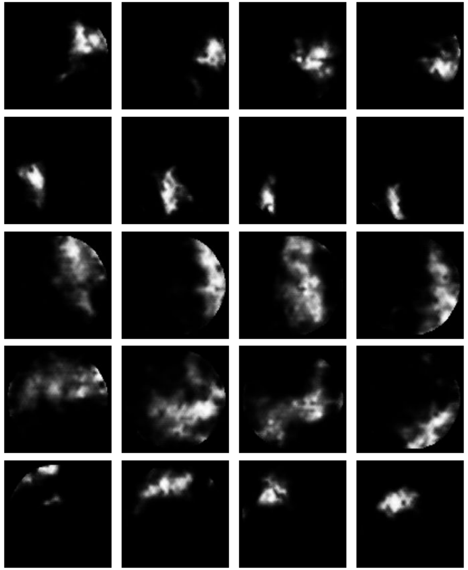


Training progress

Conditional Wasserstein GAN (CwGAN): Our Chosen Framework for Learning image-level shower features

CwGAN is a variant of classical GANs, which can be used for parameter-guided image generation

Physically relevant parameters
Θ = [size, Hillas Params] → CwGAN → Stereo-Images | Θ

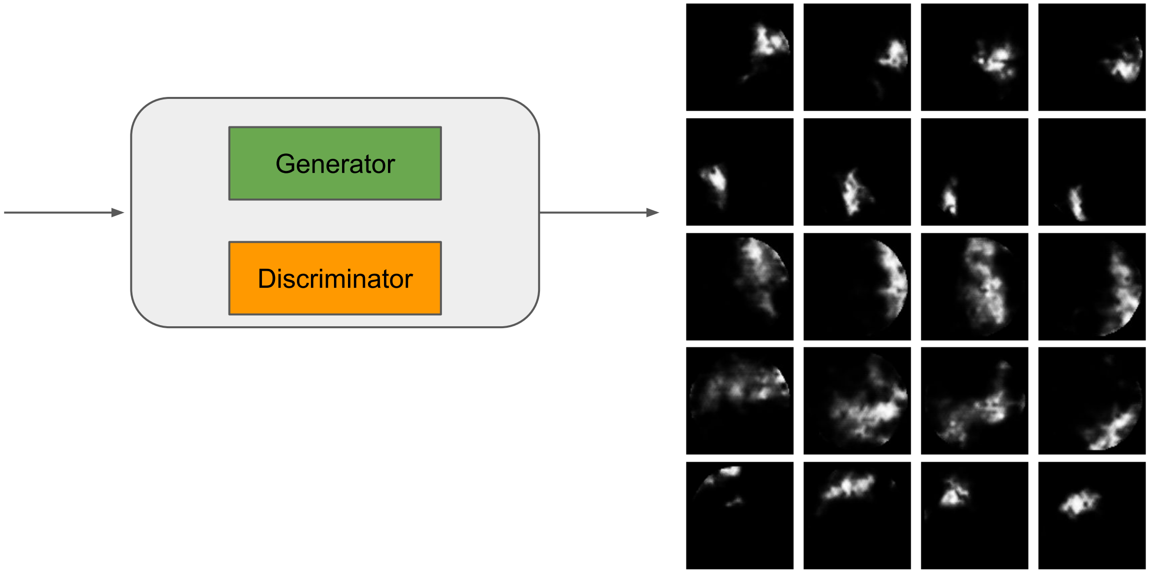


Conditional Wasserstein GAN (CwGAN): Our Chosen Framework for Learning image-level shower features

The main components of a CwGAN are a **Generator** & **Discriminator** that are trained jointly

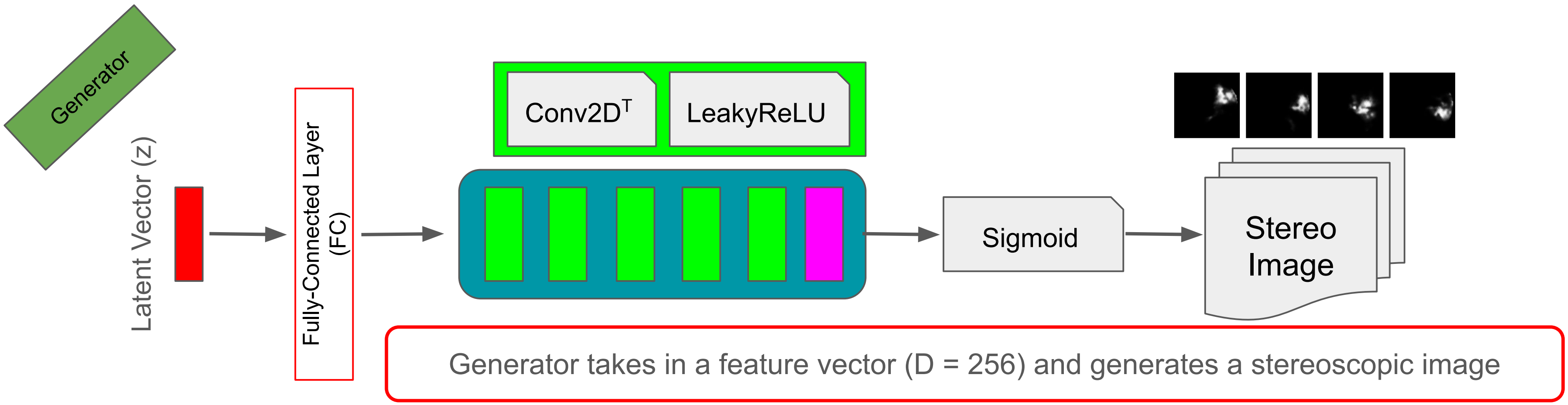
Physically relevant parameters
Θ = [size, Hillas Params]

Stereo-Images | Θ

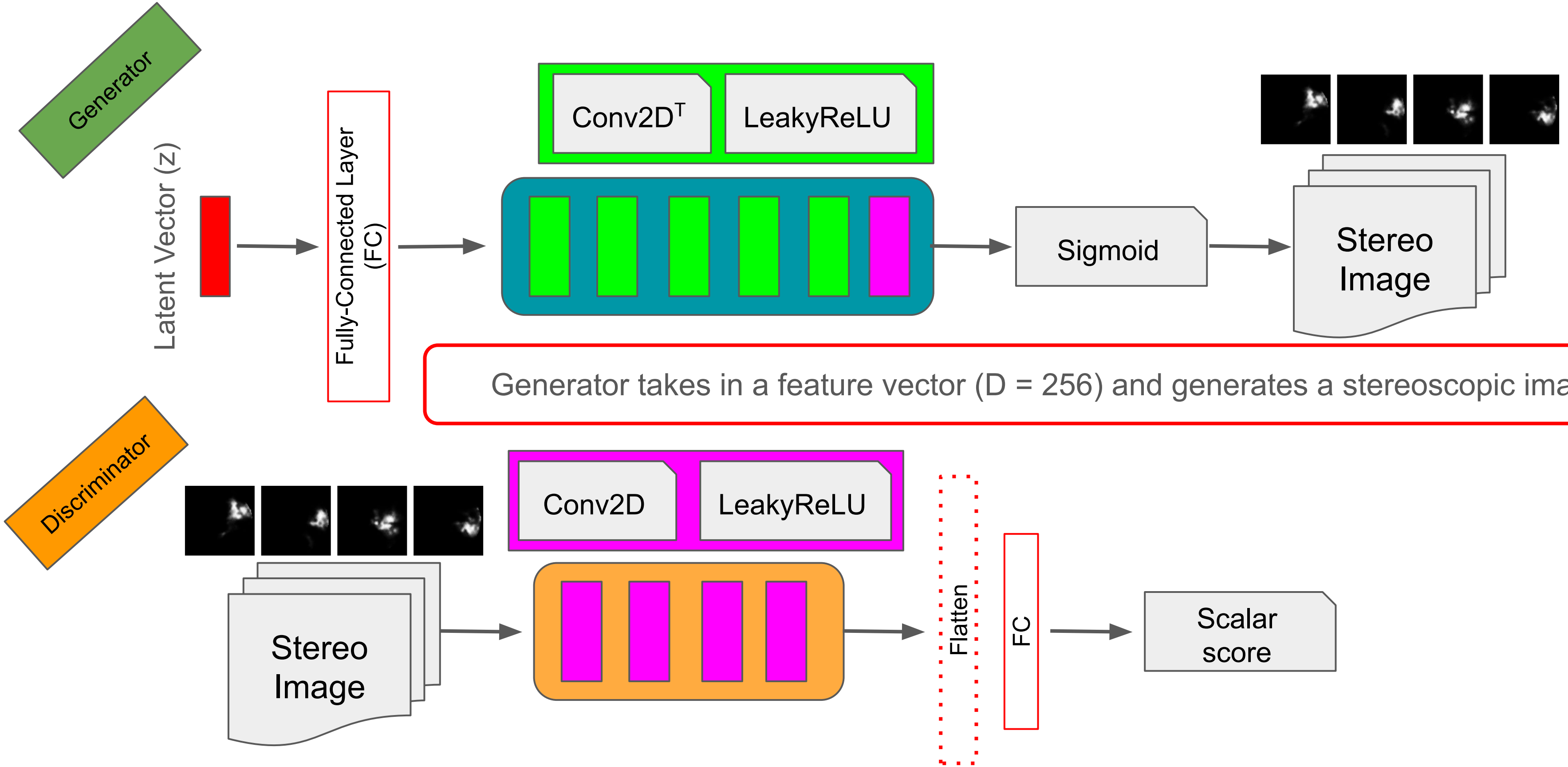


Let's unpack what CwGAN means in the next couple of slides!

Conditional Wasserstein GAN (CwGAN): Our Chosen Framework for Learning image-level shower features



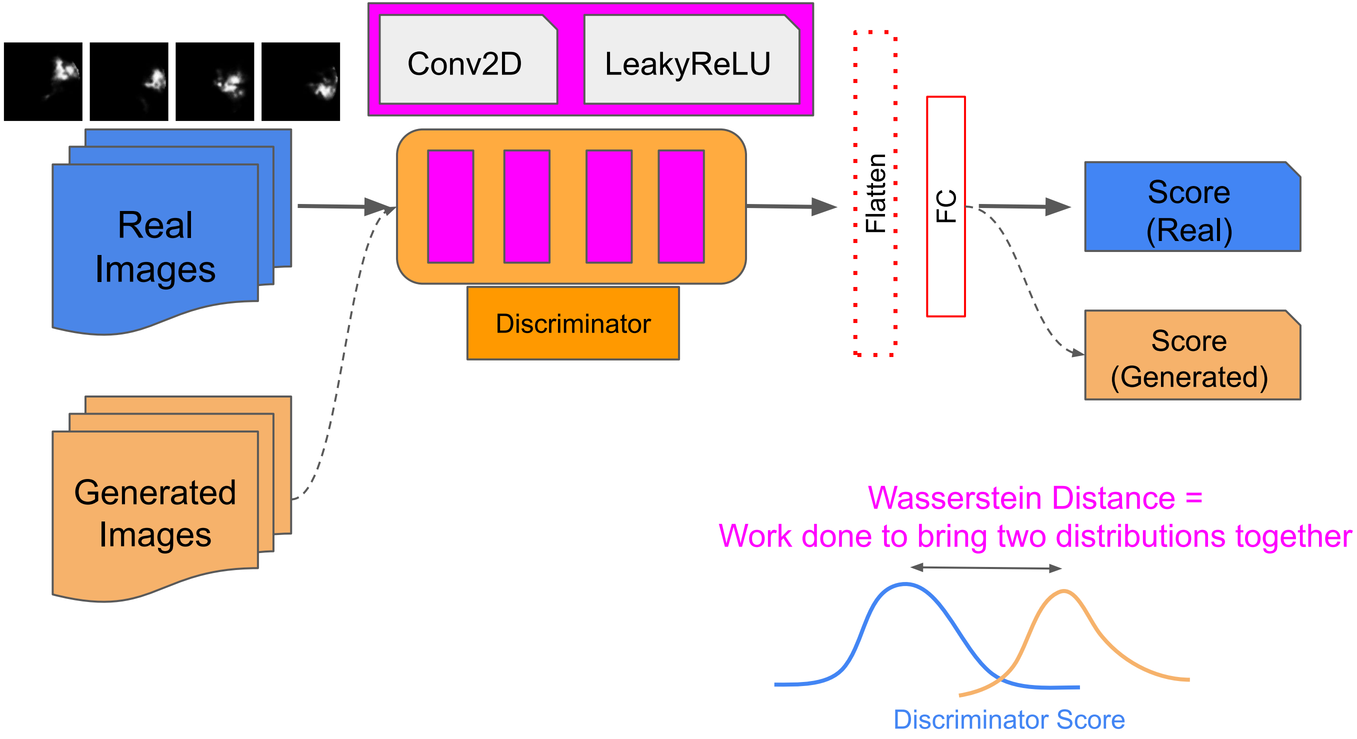
Conditional Wasserstein GAN (CwGAN): Our Chosen Framework for Learning image-level shower features



Generator takes in a feature vector (D = 256) and generates a stereoscopic image

Discriminator takes in a stereoscopic image and outputs an unbounded scalar value

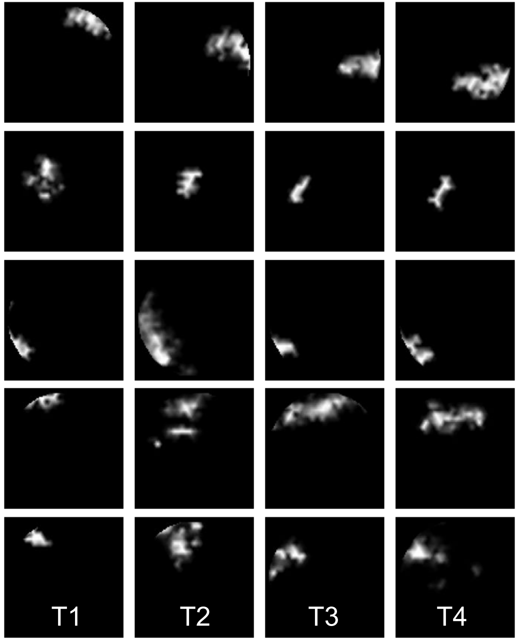
Conditional Wasserstein GAN (CwGAN): Our Chosen Framework for Learning image-level shower features

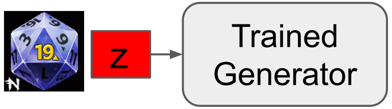


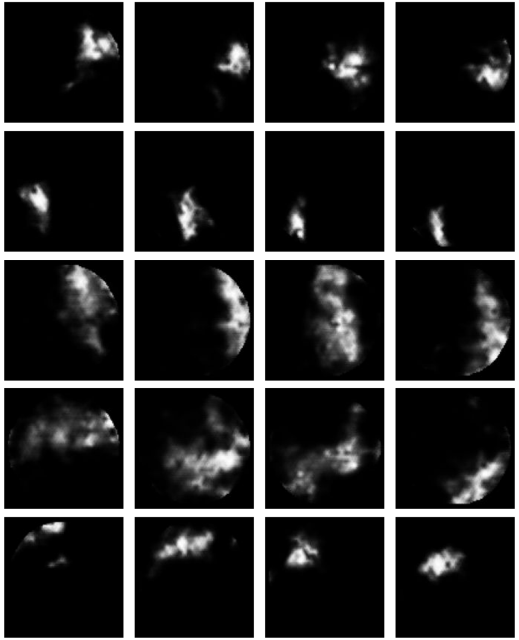
The training objective of a wGAN is to drive the **Real** and **Generated** distributions to be similar

We were able to successfully train a wGAN (unconditional) to generate stereoscopic shower images

5 Randomly Selected Real Images

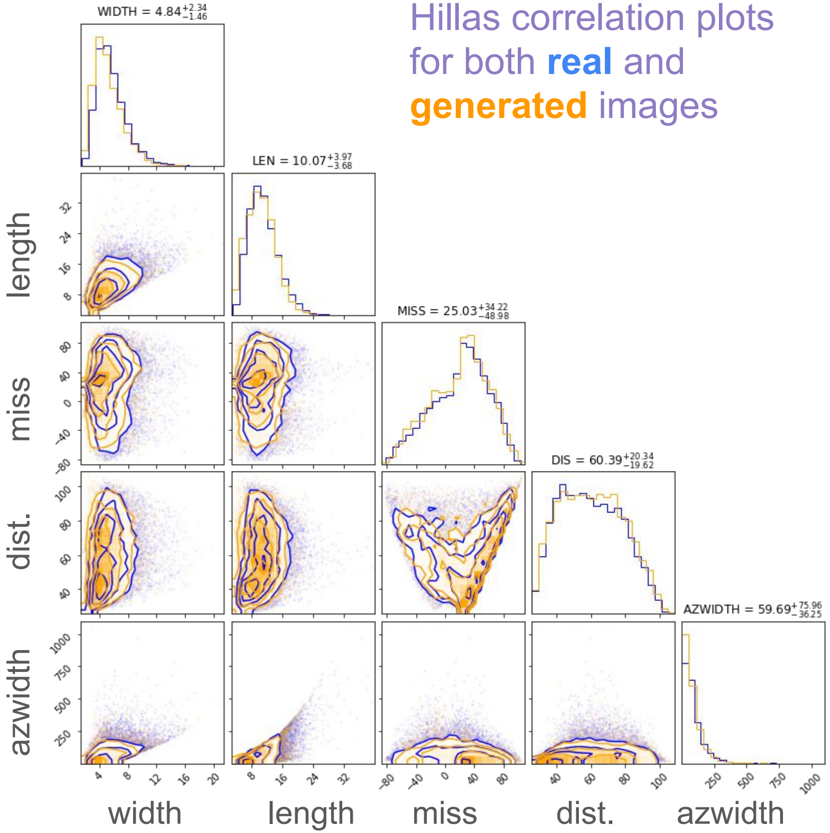




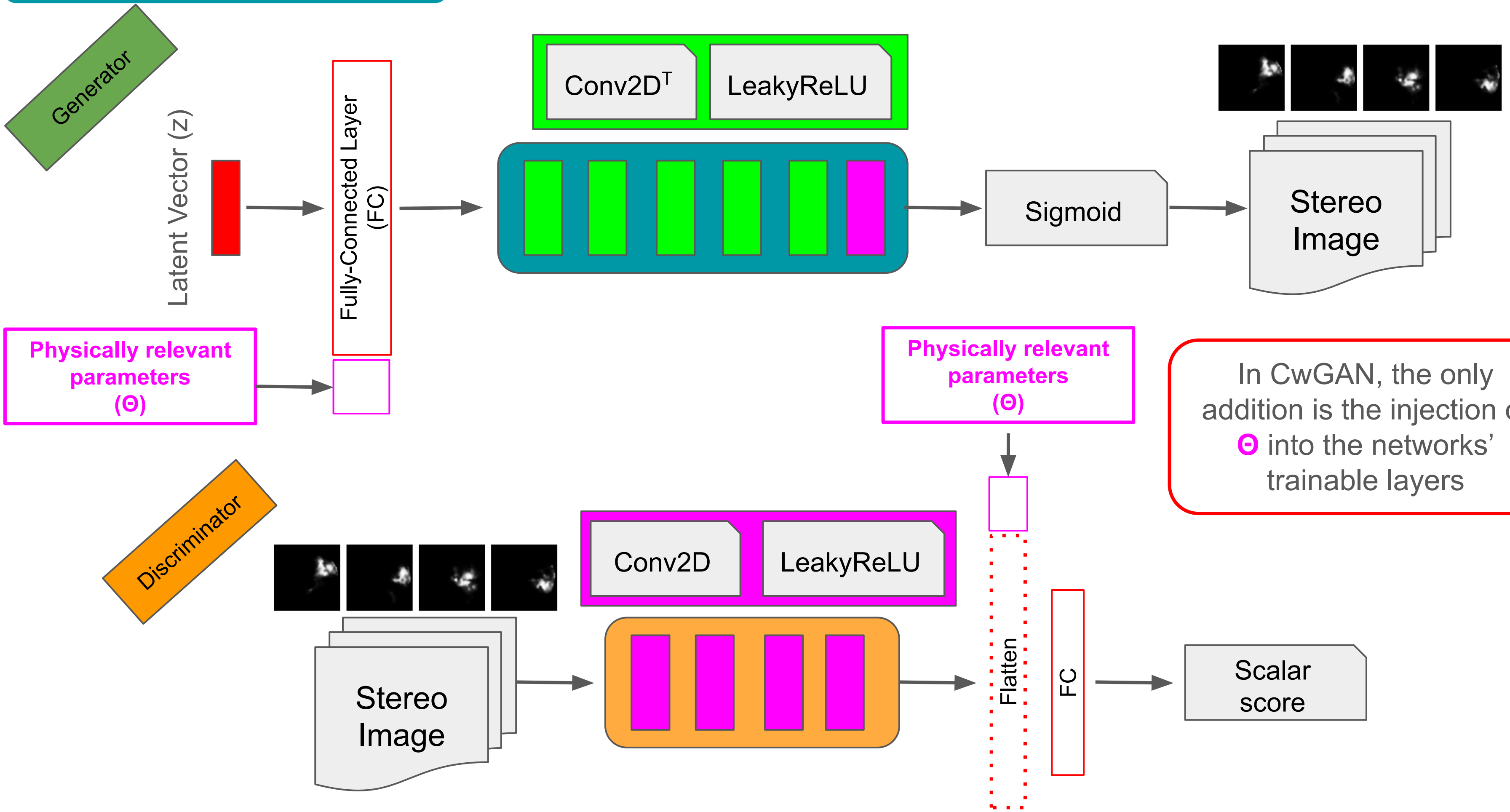


5 Randomly Generated Images

Hillas correlation plots for both **real** and **generated** images

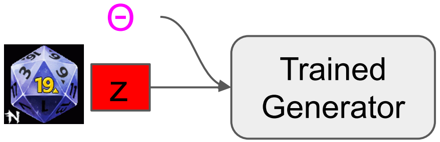


Conditional Wasserstein GAN (CwGAN): Our Chosen Framework for Learning image-level shower features

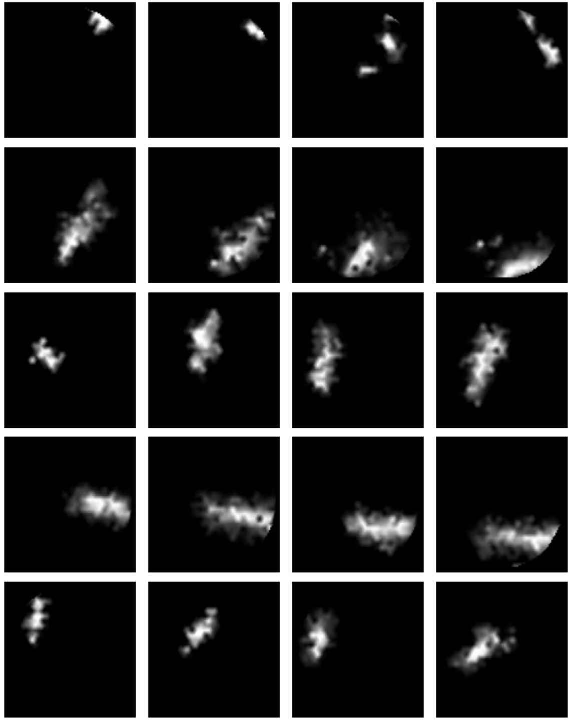


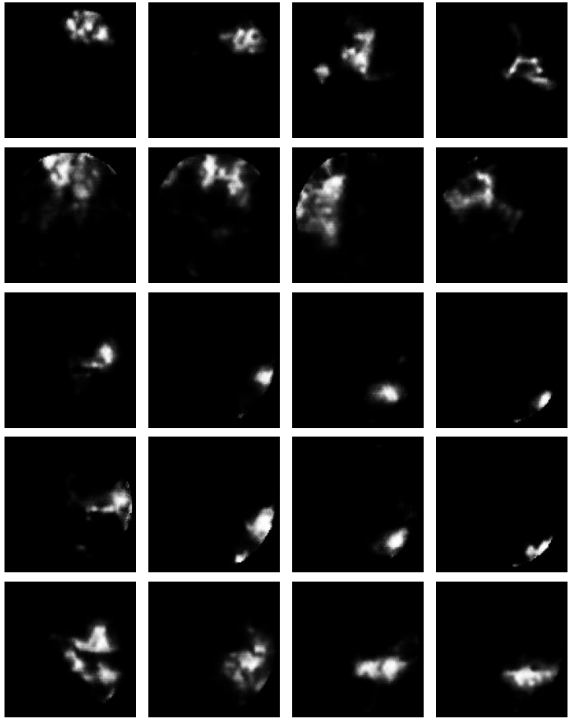
In CwGAN, the only addition is the injection of **Θ** into the networks' trainable layers

Our CwGAN model generates qualitatively promising stereoscopic shower images



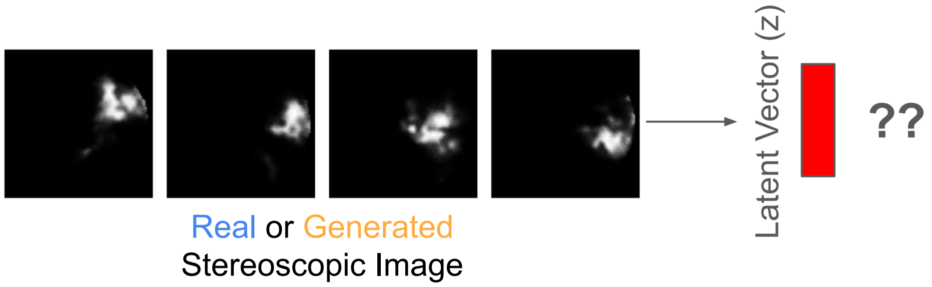
5 Randomly Chosen Real Images





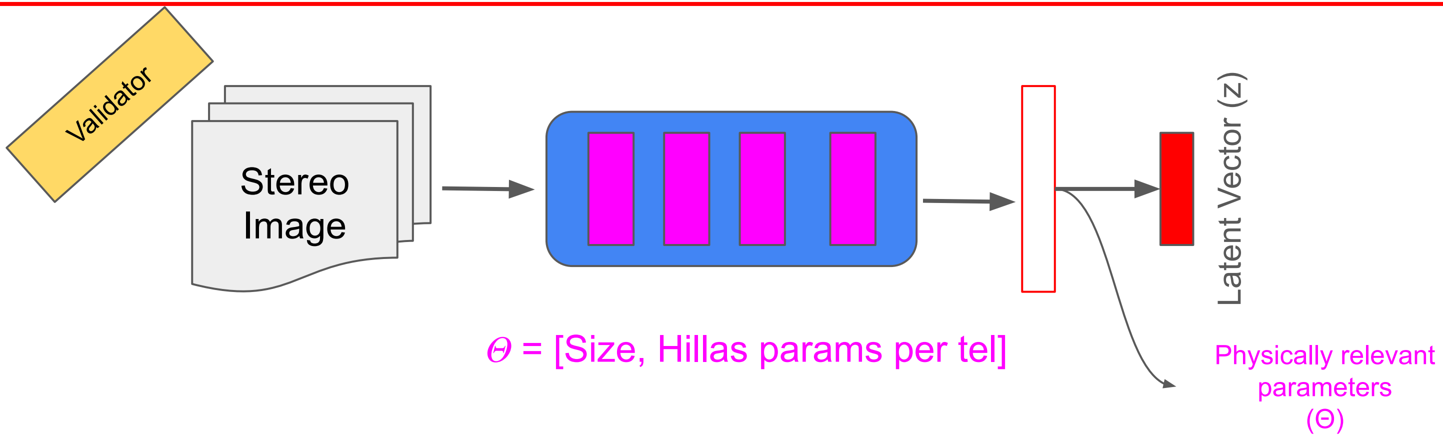
Random Generated Images conditioned on Θ

An Intermediate Validation Challenge: How to map from a given image to feature vector & parameters?

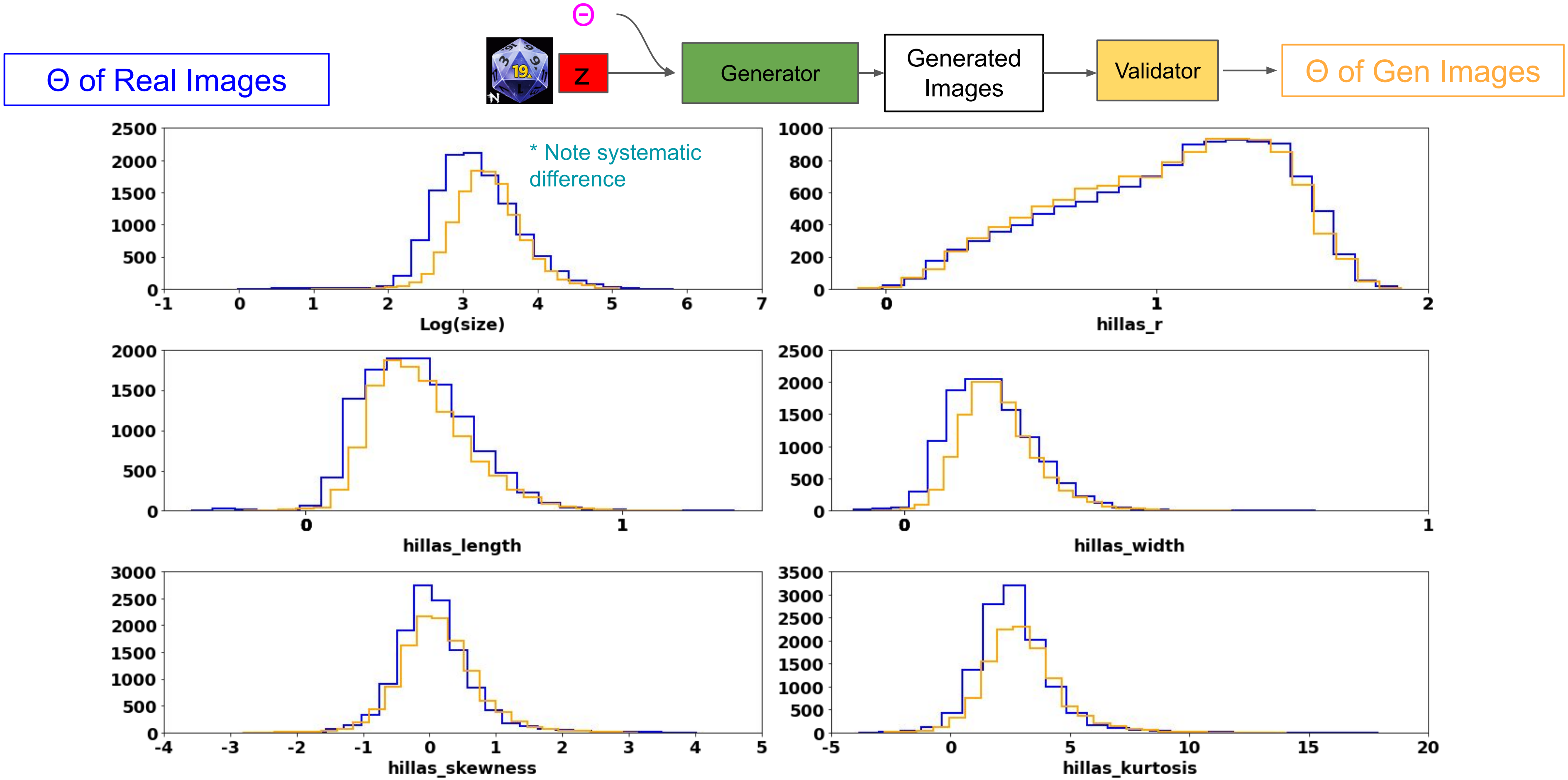


Given an image, our CwGAN framework is not equipped to map onto its corresponding feature space.

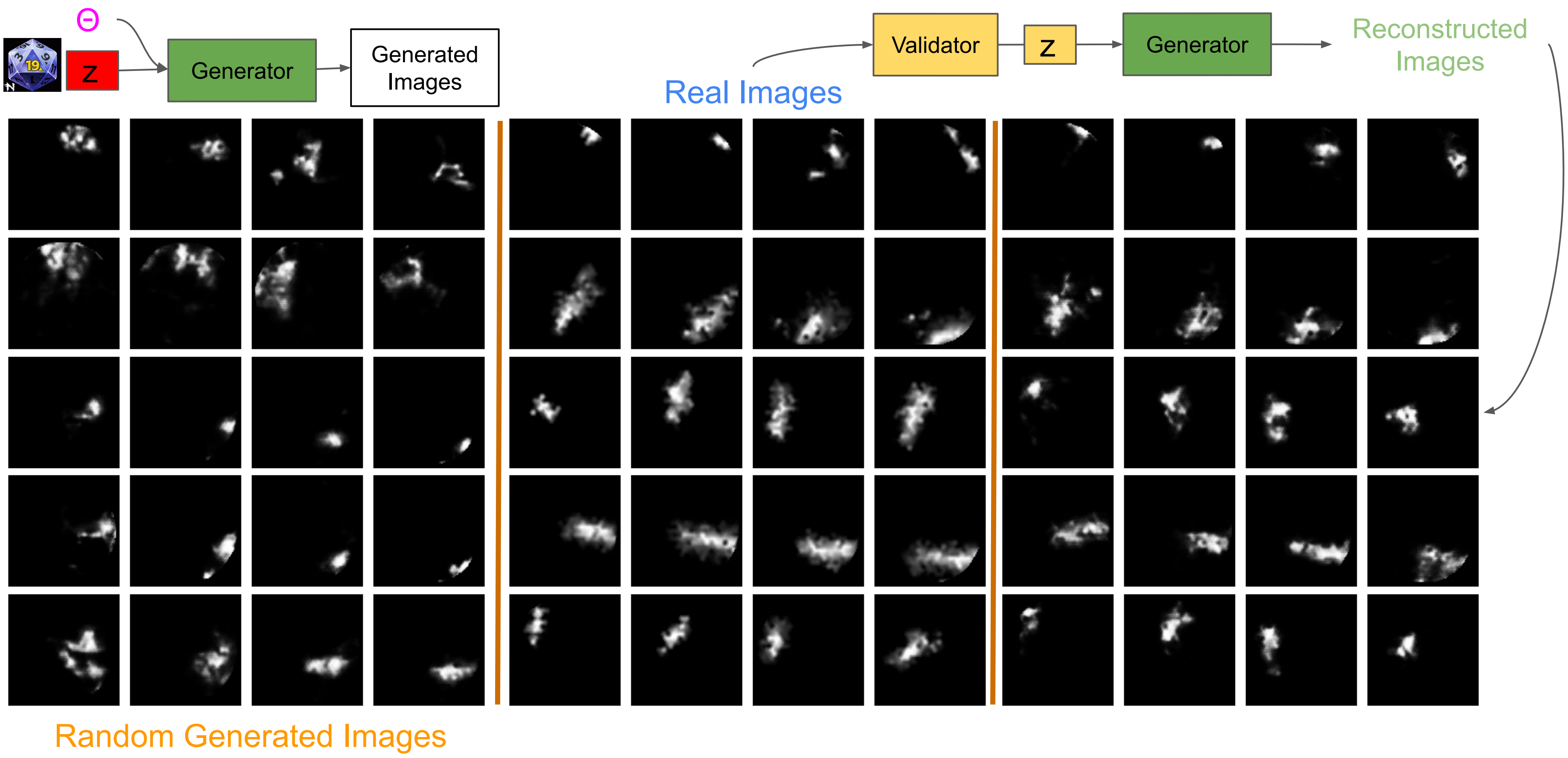
1. So, we trained a Validator model such that it can predict z from a stereoscopic image.
2. Simultaneously, we also require it to perform regression on Θ.
3. **Validator model** is trained by holding the **Generator** and **Discriminator** fixed!



Our CwGAN model-generated images demonstrate quantitative self-consistency

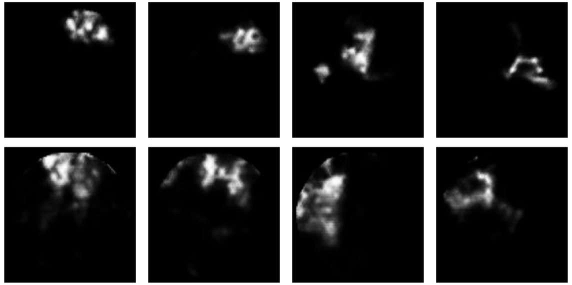


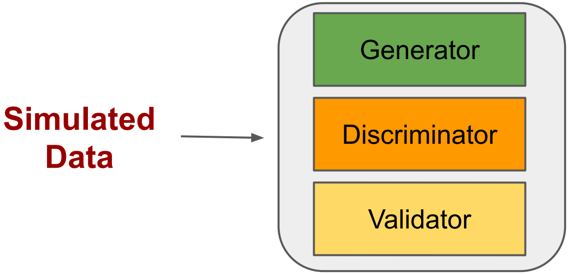
Our Validation model is OK, but could be better! – It will be when trained on more balanced data



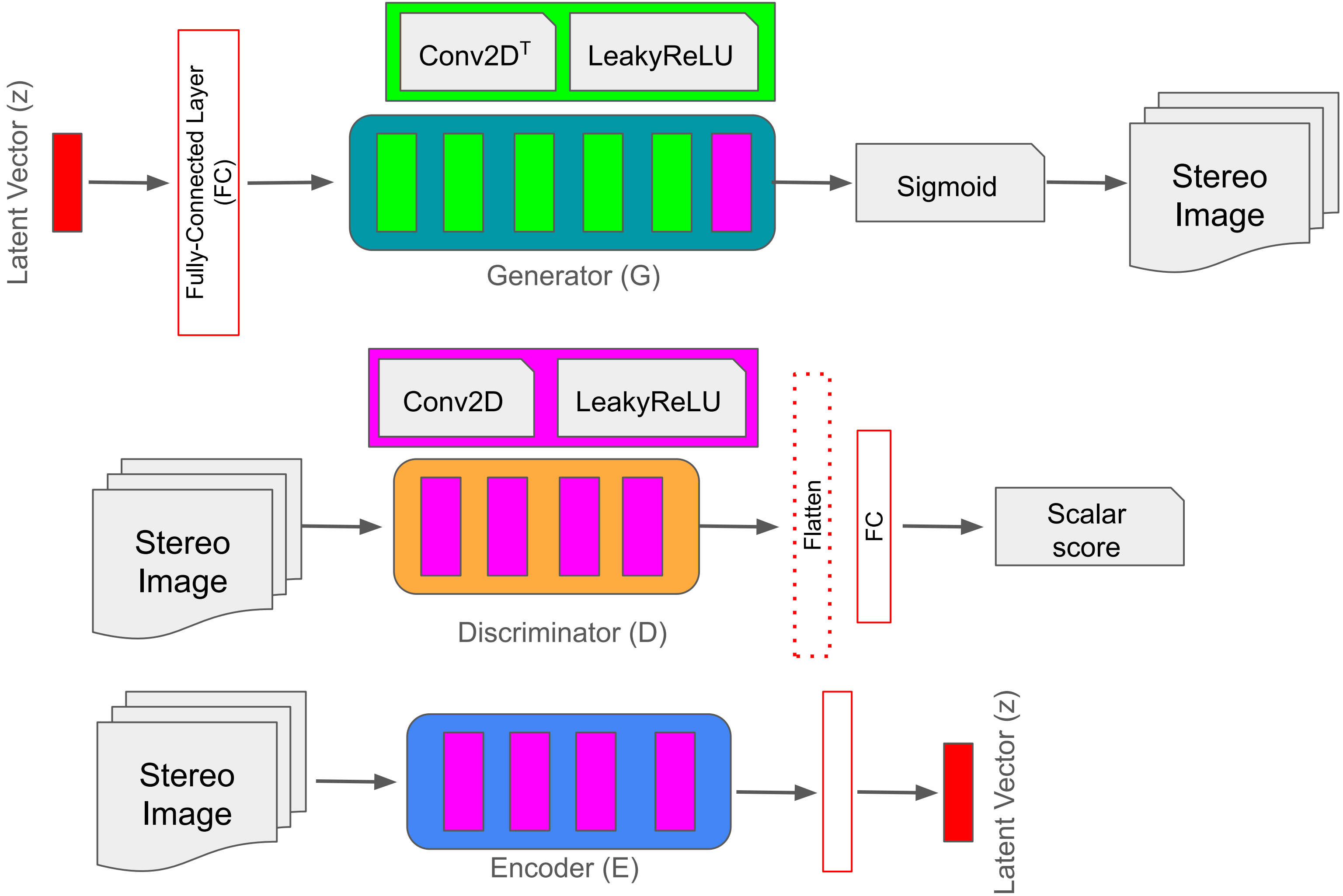
Conclusions

1. We have successfully demonstrated stereoscopic image generation with a CwGAN architecture trained on real VERITAS data
2. Next steps are to infer on simulated data to determine probabilistic gamma/hadron separation and parameter regression
3. Could also use architecture to train on simulations for fast stereoscopic image generation.

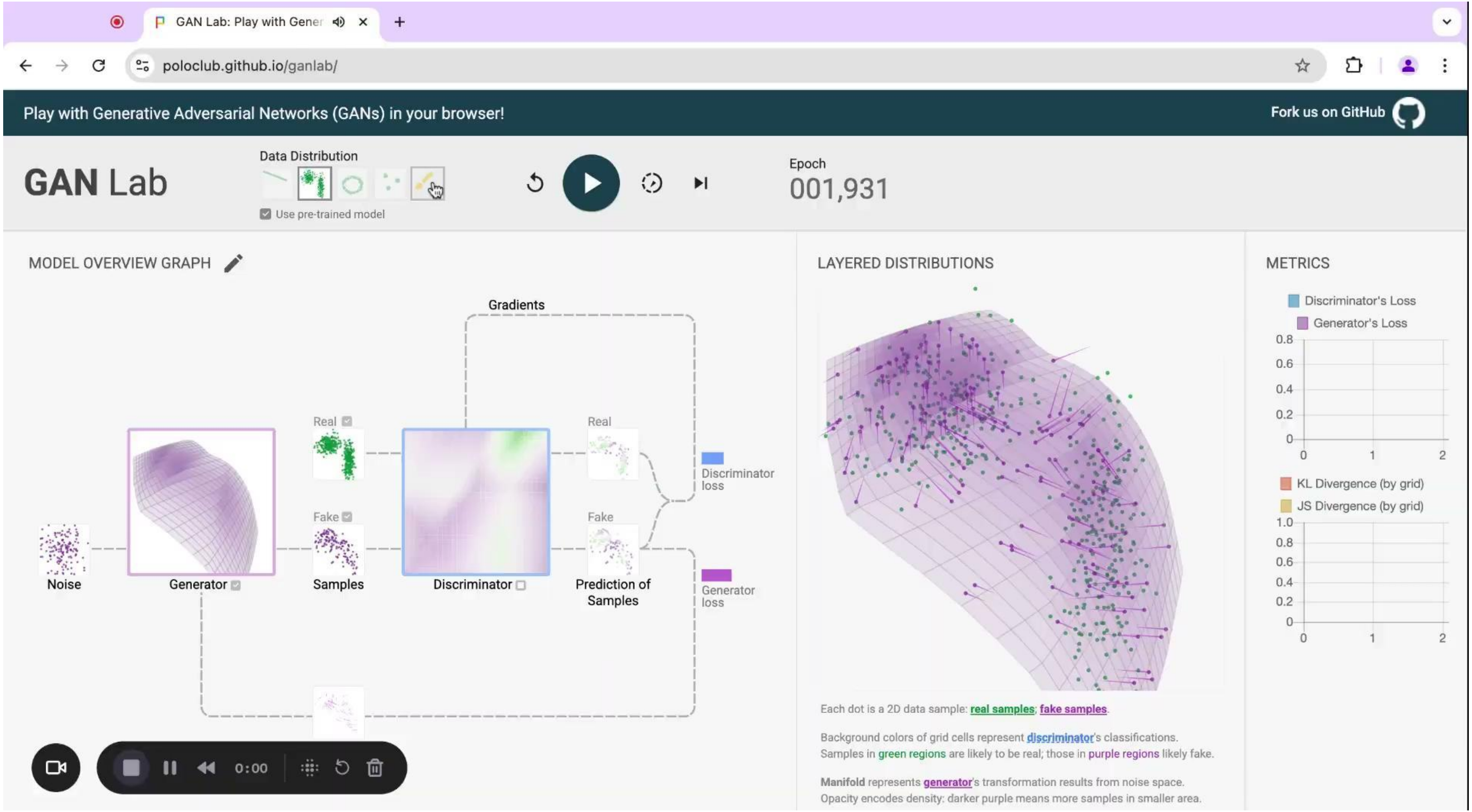




Backup Slides

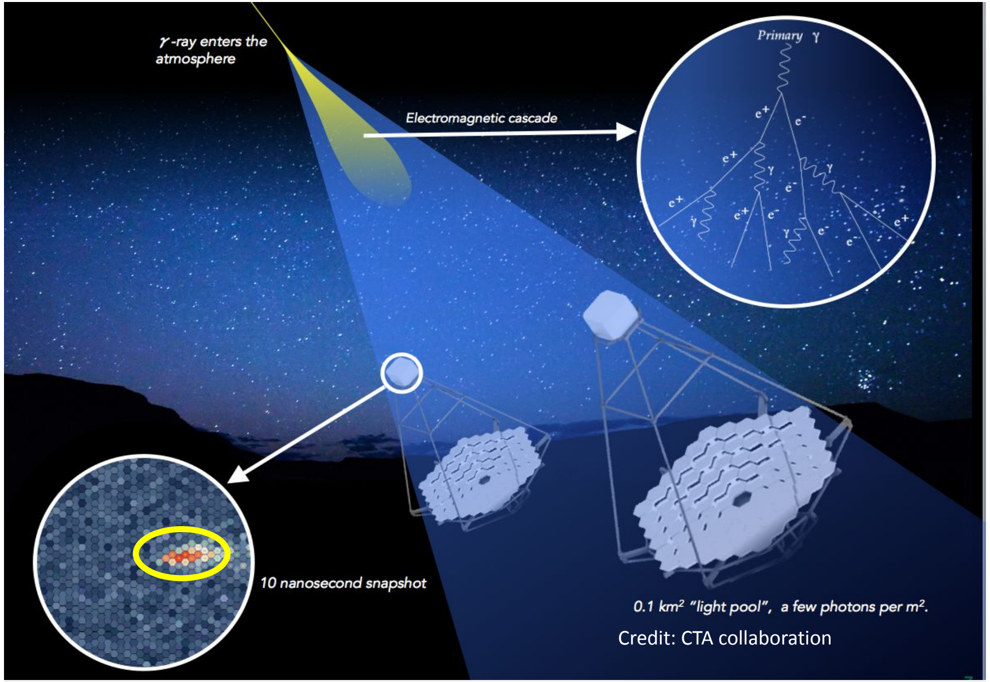


Latent Vector (z)
Fully-Connected Layer (FC)
Conv2D^T
LeakyReLU
Generator (G)
Sigmoid
Stereo Image
Conv2D
LeakyReLU
Stereo Image
Discriminator (D)
Flatten
FC
Scalar score
Stereo Image
Encoder (E)
Latent Vector (z)



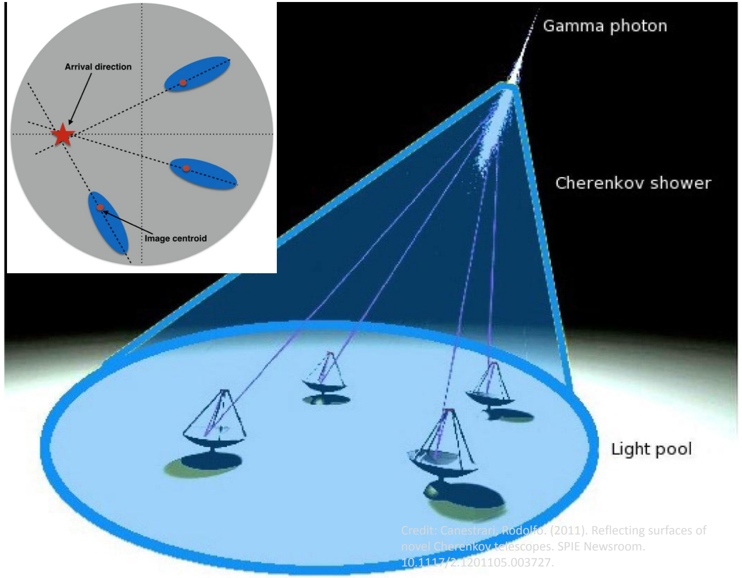
GAN Lab: Play with Gener
poloclub.github.io/ganlab/
Play with Generative Adversarial Networks (GANs) in your browser!
Fork us on GitHub
GAN Lab
Data Distribution
Use pre-trained model
Epoch
001,931
MODEL OVERVIEW GRAPH
Gradients
Real
Fake
Noise
Generator
Samples
Discriminator
Prediction of Samples
Discriminator loss
Generator loss
0:00
LAYERED DISTRIBUTIONS
Each dot is a 2D data sample: real samples; fake samples.
Background colors of grid cells represent discriminator's classifications.
Samples in green regions are likely to be real; those in purple regions likely fake.
Manifold represents generator's transformation results from noise space.
Opacity encodes density: darker purple means more samples in smaller area.
METRICS
Discriminator's Loss
Generator's Loss
KL Divergence (by grid)
JS Divergence (by grid)

Imaging Atmospheric Cherenkov Telescope (IACT)



Credit: CTA collaboration

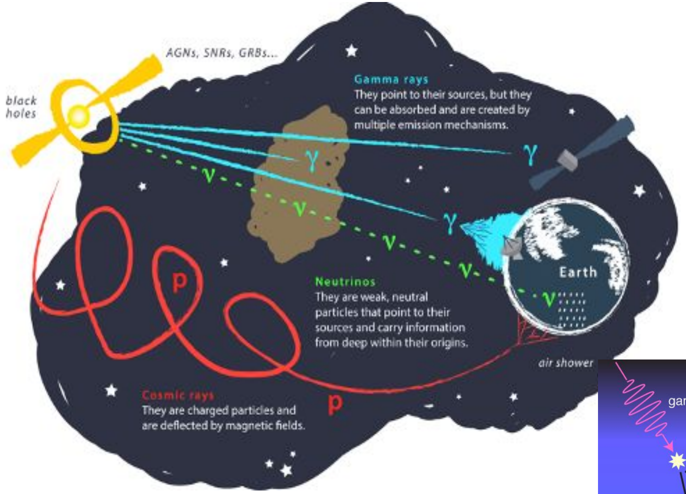
Cherenkov radiation emitted by the cascade of secondary particles is captured using Photomultiplier Tubes. Properties of the primary are reconstructed with the help of Monte Carlo simulations of EAS and the corresponding Cherenkov image.



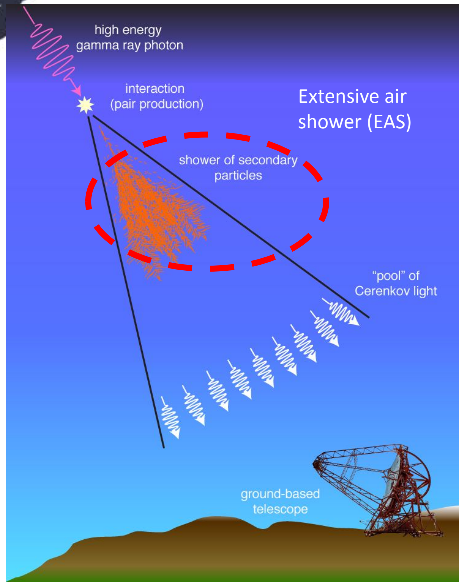
Credit: Canestrari, Rodolfo. (2011). Reflecting surfaces of novel Cherenkov telescopes. SPIE Newsroom. 10.1117/2.1201105.003727.

Stereoscopic imaging:
1. Better reconstruction of direction, core location, and energy of primary with multiple telescopes
2. Elimination of fluctuations in low energy range due to night sky background and muons by applying coincident trigger criteria

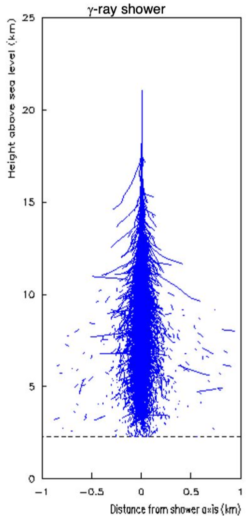
Gamma-Hadron Separation: A Critical Scientific Analysis Step

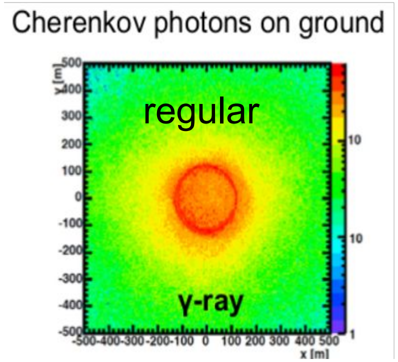


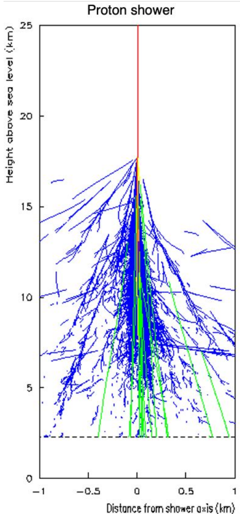
Both gamma-rays and cosmic rays (protons) produce EAS on entering the atmosphere.

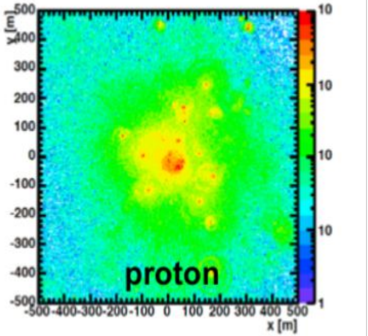


There are about 1000-10000 cosmic-ray-initiated showers for every gamma-ray initiated one.







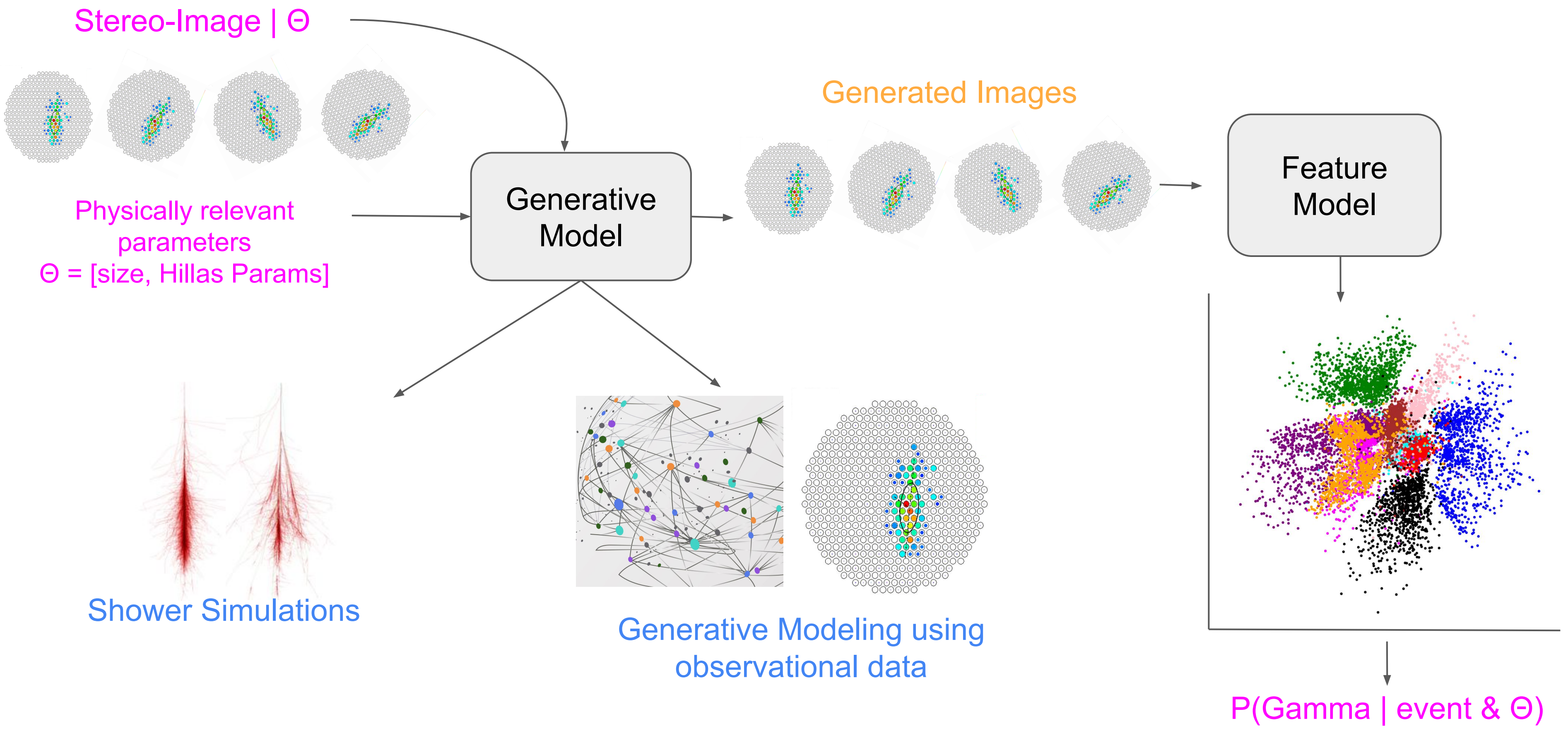


VERITAS

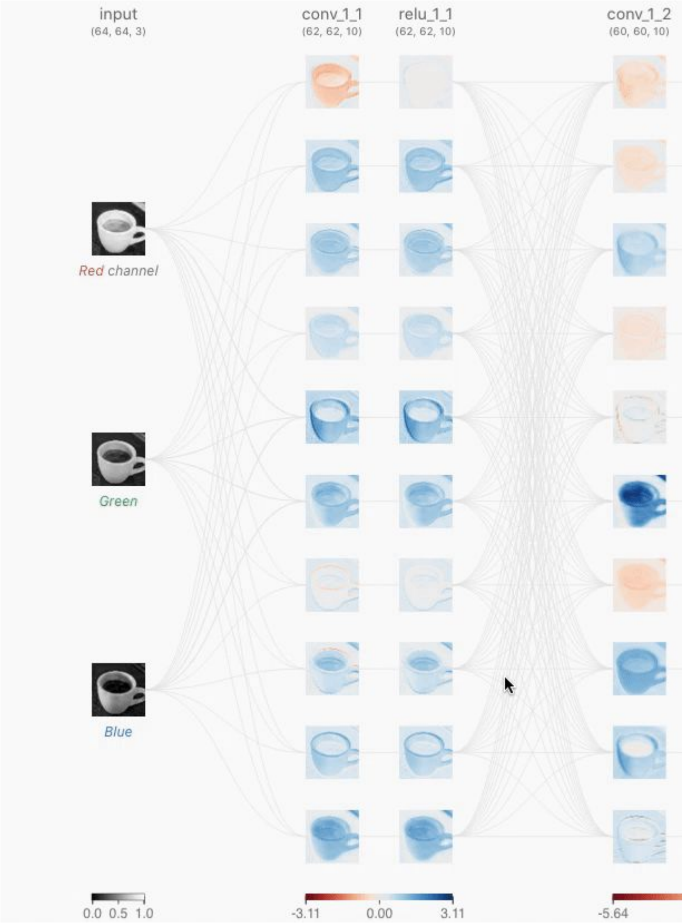
The Very Energetic Radiation Imaging Telescope Array System

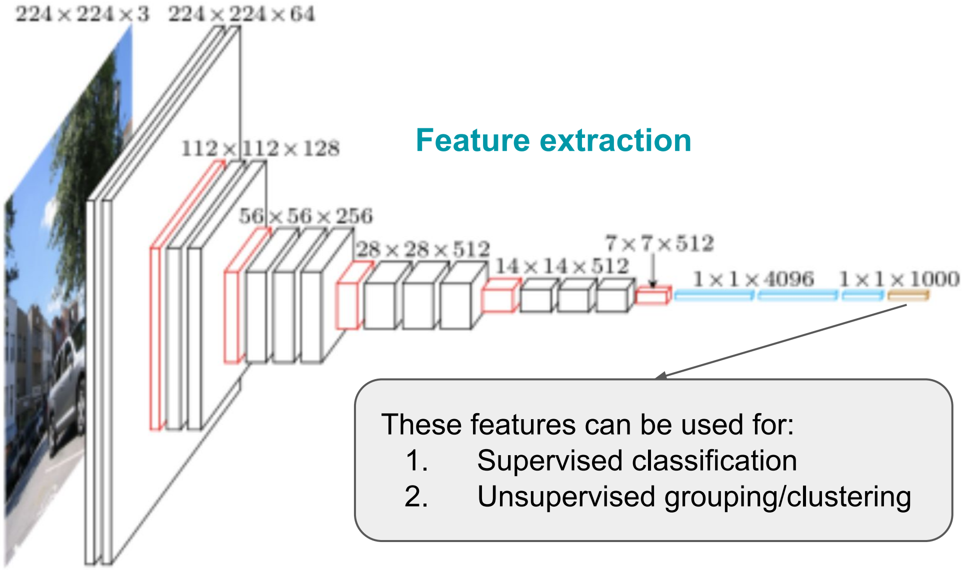
- An array of four 12m-diameter imaging atmospheric Cherenkov telescopes
- Located at the Fred Lawrence Whipple Observatory in southern Arizona
- Energy range: 85 GeV to >30 TeV — Energy resolution: 17%
- Angular resolution: ~0.08 @ 1 TeV — Source location accuracy: error < 50 arcsec
- Sensitivity: 1% Crab in ~25h

Deep Generative Learning as an avenue for data-driven learning of shower-image features



Convolutional Neural Networks (CNNs) as feature extraction modules





Feature extraction

These features can be used for:

1. Supervised classification
2. Unsupervised grouping/clustering