TEST RESULTS AND SUMMARIES
EVALUATION TEST: DEEPFALCON
PROJECT: “Graph Representation Learning for Fast Detector Simulation”
CONDUCTED BY- RUSHIL SINGHA
ORGANIZATION: ML4SCI-DEEPFALCON

Q2)-> Common Task 2. Jets as graphs

- Please choose a graph-based GNN model of your choice to classify (quark/gluon) jets. Proceed as follows:

1. Convert the images into a point cloud dataset by only considering the non-zero pixels for every event.
2. Cast the point cloud data into a graph representation by coming up with suitable representations for nodes and edges.
3. Train your model on the obtained graph representations of the jet events.

Discuss the resulting performance of the chosen architecture.

IMPLEMENTATION->

Model Architecture->

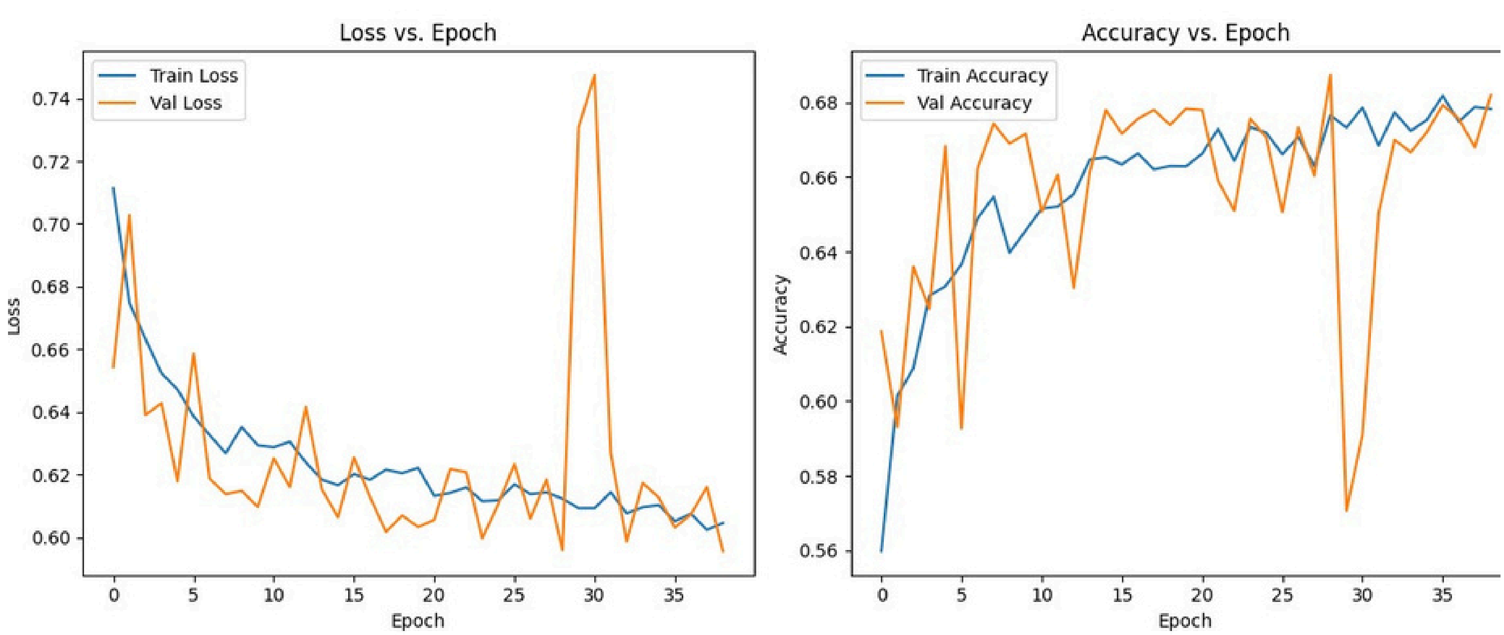
1) Designed a hybrid GNN model that combines three different graph convolution types:
   This architecture:

   a) Uses **GAT** (Graph Attention Networks) for the first layer, which can leverage edge features and attention mechanisms
   b) Employs **GraphSAGE** as the second layer to effectively aggregate neighborhood information
   c) Includes a **GIN** (Graph Isomorphism Network) layer which is theoretically the most expressive GNN variant

Point to Consider->

1) Took only 20,000 samples out of ~130000 samples in the dataset due to gpu and time constraints but as far as GNN are concerned we know as samples size increases the accuracy and loss gets better

2) But even with 20,000 sample the graph was able to give out these->
   TEST RESULTS->
   TEST ACCURACY– ~68%
   ROC AUC SCORE– 0.7339



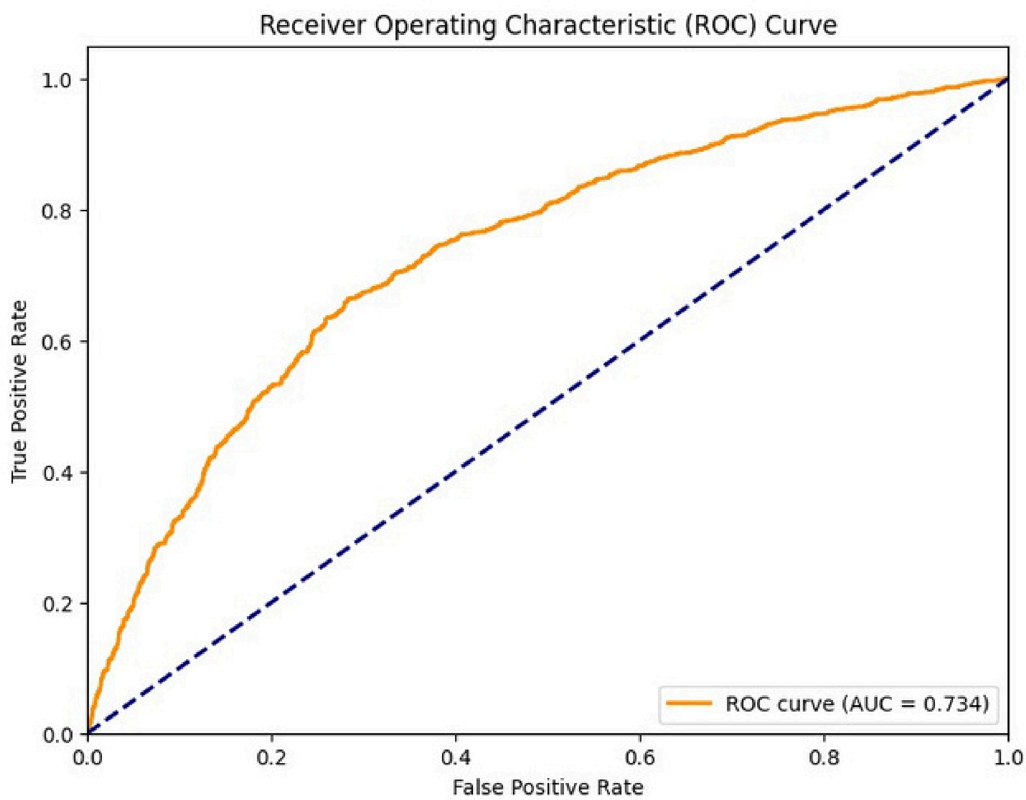
FINDINGS->

1) Loss vs Epochs->

Both training and validation loss gradually decreased over epochs, with the expected fluctuations in validation metrics, however a notable validation loss spike occurred around epoch 30, but the model recovered Accuracy for both training and validation sets steadily improved, plateauing around 67-68%

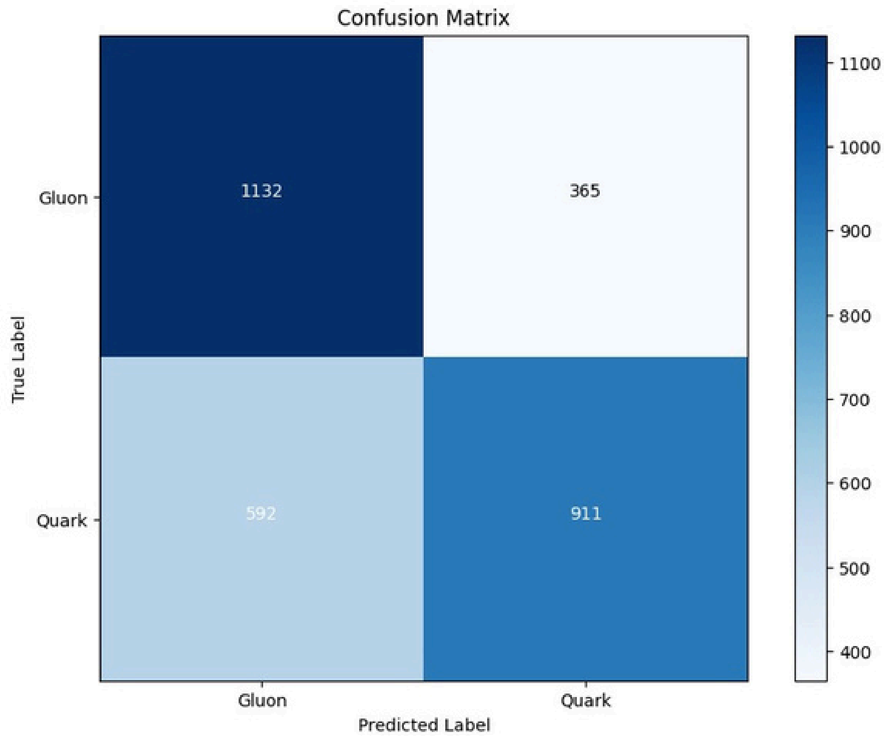
2) Accuracy vs Epochs->

Similarly for accuracy as expected it gradually increases with Epochs, but there is a drop during 30th epoch similar to what we saw in loss vs epoch graph but that might be due to some data leak or outlier in the data, the accuracy ends up plateauing around 68% accuracy



3) ROC Curve->

The ROC curve with AUC of 0.7339 indicates good discrimination ability, with more inclement towards true positive rate than false positive rate and it performing significantly better than random classification, though there's still room for improvement compared to an ideal classifier, which can be easily covered if trained on a sample set of bigger size.



4) Analysis of the Confusion Matrix->

True Gluons: 1132 correctly classified as gluons, 365 misclassified as quarks True Quarks: 911 correctly classified as quarks, 592 misclassified as gluons The model shows better performance at identifying gluons (higher true positive rate) than quarks.

Final Summary->

I implemented a graph-based hybrid-GNN model for quark/gluon jet classification that achieved 68.10% accuracy and an AUC score of 0.7339. The model performs particularly well at identifying gluons, with a strong true positive rate as shown in the confusion matrix. Despite training on only 20,000 of the available 130,000 samples, the model showed consistent improvement over 39 epochs with promising discrimination ability between jet types.

Q1)-> Common Task 1. Auto-encoder of the quark/gluon events

- Please train a variational auto-encoder to learn the representation based on three image channels (ECAL, HCAL and Tracks) for the dataset.
- Please show a side-by-side comparison of original and reconstructed events.

IMPLEMENTATION->
I trained the variational autoencoder for 200 epochs due to the complex patterns of jet sprays. When I used fewer epochs and lower latent dimensions, the model struggled to detect high-energy deposits. Increasing the latent dimension and training duration improved the results, yielding reasonable outputs when I used approximately 10,000 samples. Here are the results:

Final Metrics

- MSE: 0.000224: very low, indicating good overall reconstruction
- MAE: 0.002218: low absolute error
- RMSE: 0.014955 : square root of MSE
- NRMSE: 0.014955: normalized RMSE)
- PSNR: 36.504090 : peak signal-to-noise ratio - higher is better, indicating good quality

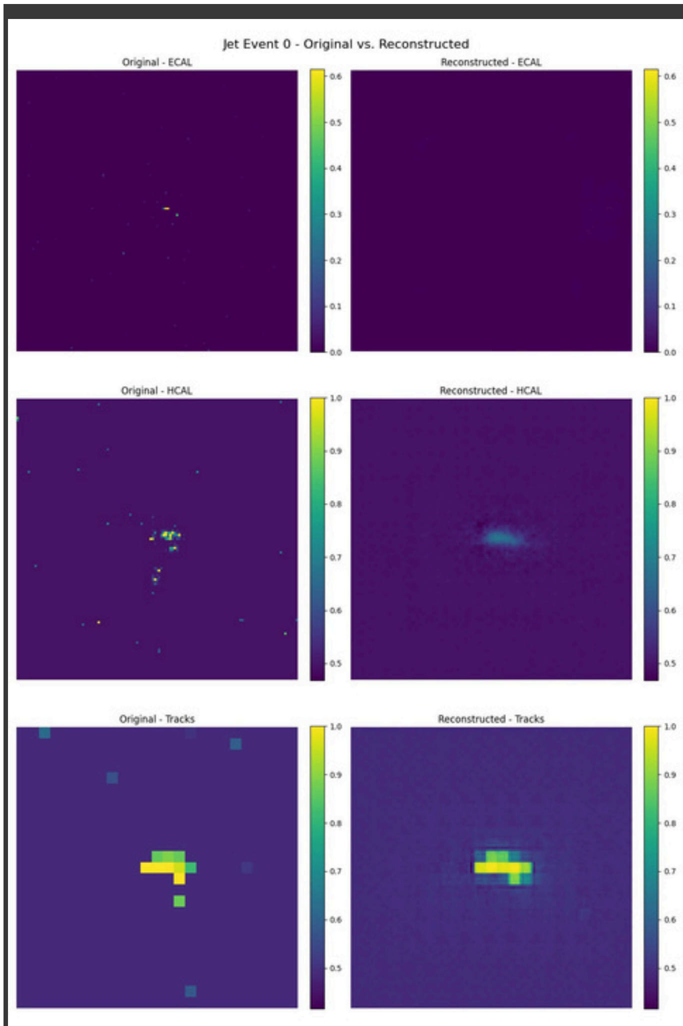
I AM VISUALIZING THE RESULTS OF 3 SAMPLES DURING TRAINING PROCESS

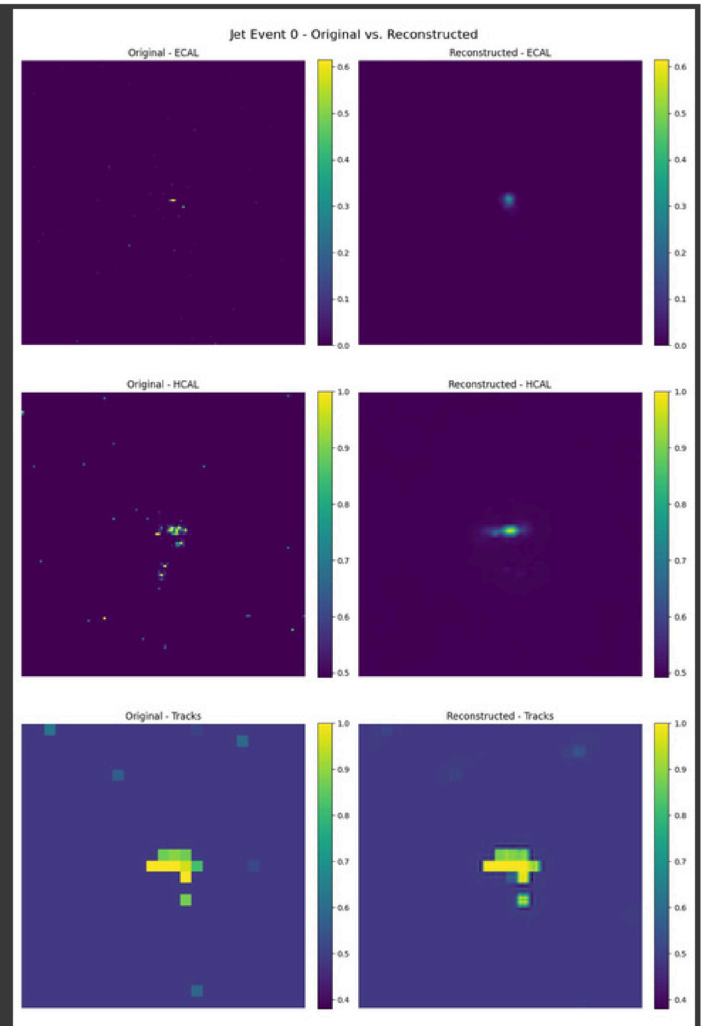
1) THE FIRST IMAGE COMPARES THE ECAL, HCAL AND TRACKS OF ORIGINAL VS RECONSTRUCTED IMAGES WHEN TRAINING PROCESS JUST STARTED(INITIAL EPOCH) AND
2) SECOND IMAGE IN THE SAME ROW COMPARES THE ECAL,HCAL AND TRACK OF ORIGINAL IMAGE VS FINAL EPOCH(FINAL RESULT).

1) Sample 1->

ORIGINAL VS RECONSTRUCTION (INITIAL EPOCH)

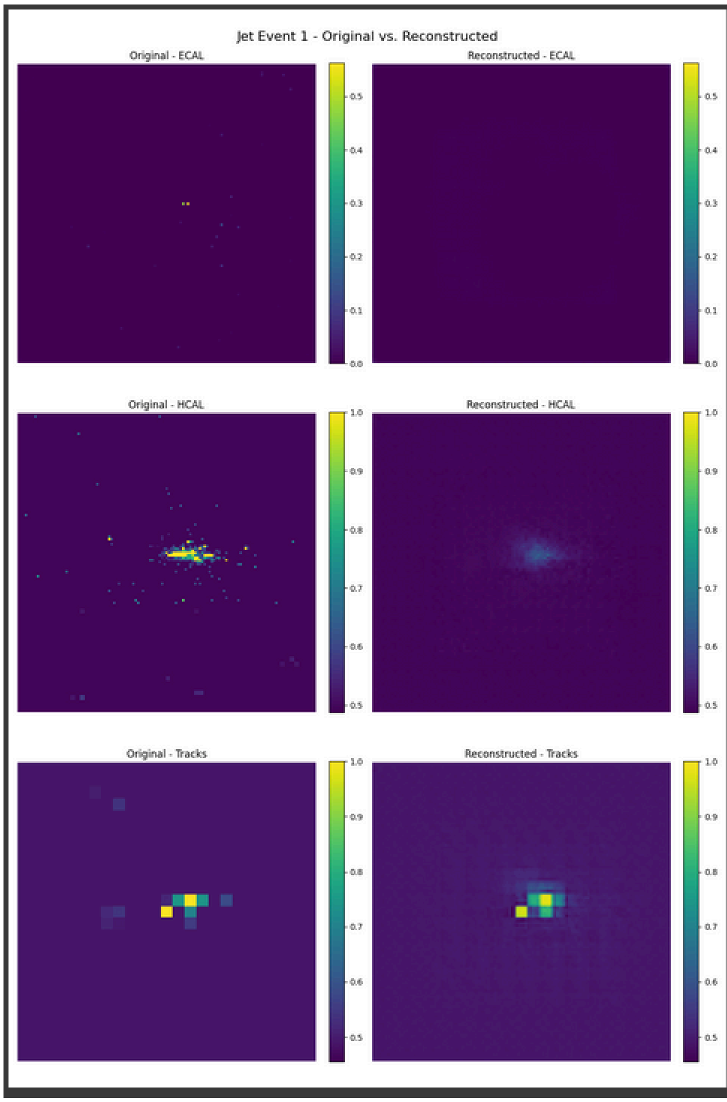
ORIGINAL VS RECONSTRUCTION (FINAL EPOCH)



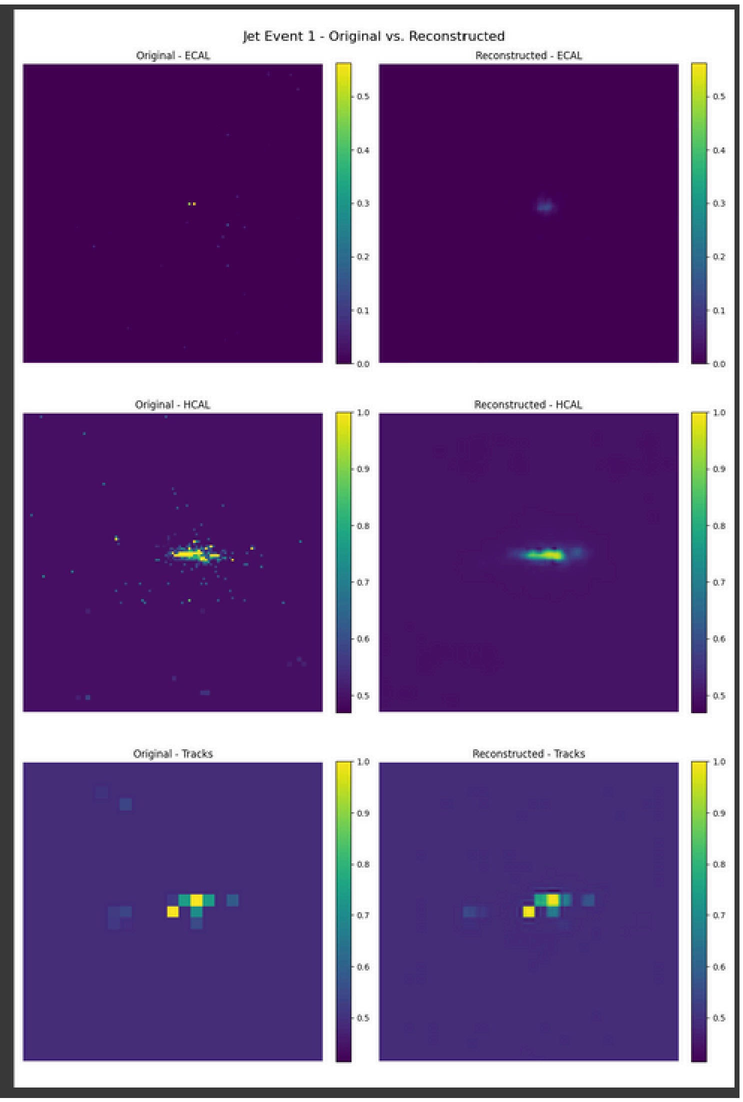


2) Sample 2->

ORIGINAL VS RECONSTRUCTION (INITIAL EPOCH)

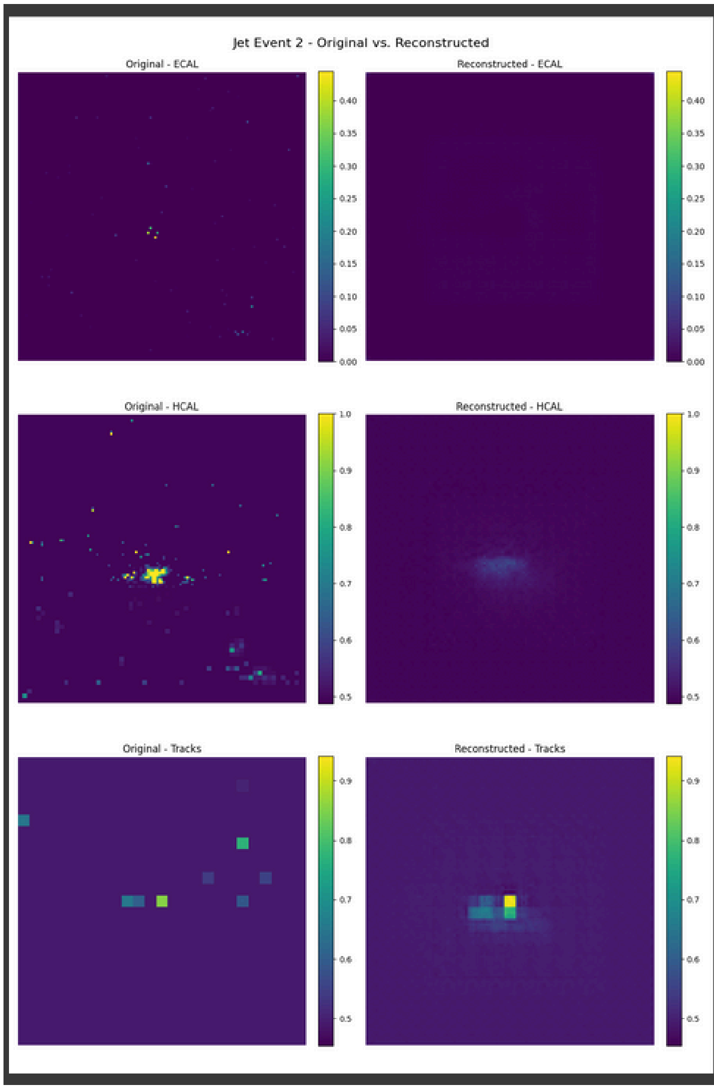


ORIGINAL VS RECONSTRUCTION (FINAL EPOCH)

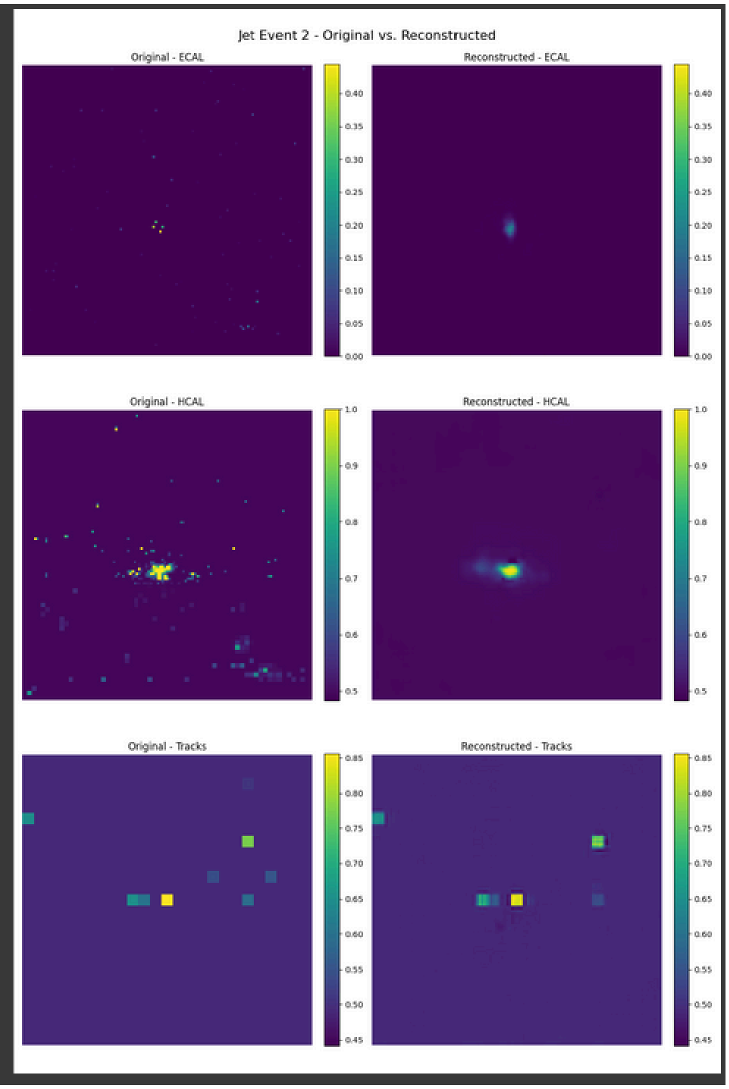


3) Sample 3->

ORIGINAL VS RECONSTRUCTION (INITIAL EPOCH)



ORIGINAL VS RECONSTRUCTION (FINAL EPOCH)



Conclusion-> **Training Progress**

Initial Epoch (Image 1):

- ECAL: Almost no reconstruction visible in the right panel
- HCAL: General jet shape captured but with significant blurring and loss of detailed structure
- Tracks: Core structure preserved but peripheral tracks poorly reconstructed

Final Epoch (Image 2):

- ECAL: Now shows visible reconstruction with proper positioning but reduced intensity
- HCAL: Much improved definition with better color intensity matching and core structure
- Tracks: Clear improvement in both core and peripheral track reconstruction

q3)- (specific task-1)->

- Please train a simple graph autoencoder on this dataset. Please show a visual side-by side comparison of the original and reconstructed events and appropriate evaluation metric of your choice. Compare to the VAE model results.

IMPLEMENTATION->

I implemented a simple graph autoencoder on the given dataset of jet sprays from quarks and gluons. To analyze the results, I visualized the original graphs alongside their reconstructed versions. Specifically, I compared five samples from the initial epoch with five samples from the final trained model to observe how the encoder improved with training and effectively reduced losses.

I followed the same metrics i followed for task 1->

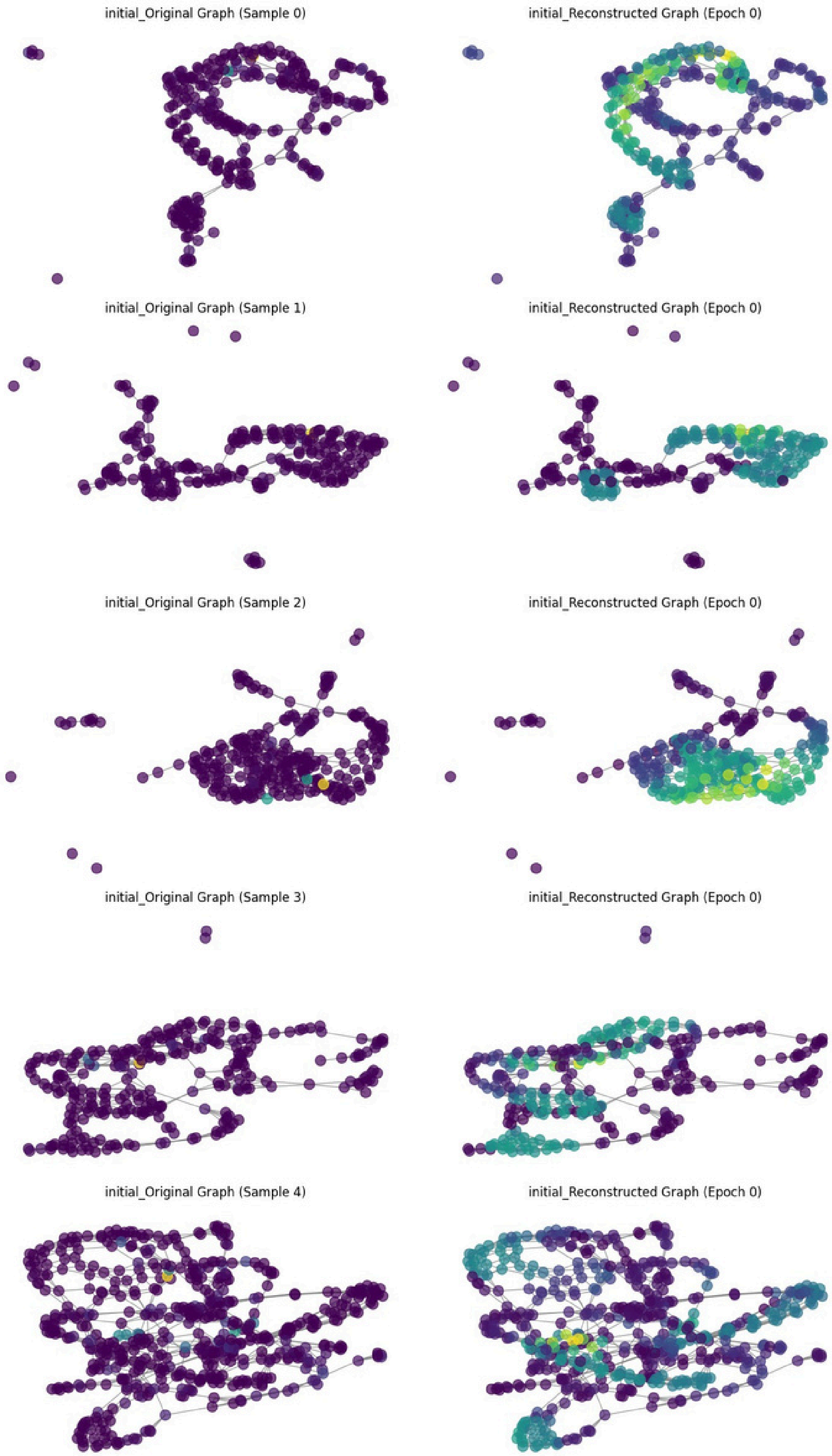
Simple Graph Auto-Encoder Metrics

The simple graph auto-encoder, has the following metrics:

- Mean Squared Error (MSE): 0.005346
- Mean Absolute Error (MAE): 0.028314
- Root Mean Square Error (RMSE): 0.073041
- Normalized Root Mean Square Error (NRMSE): 0.073041
- Peak Signal-to-Noise Ratio (PSNR): 22.737680

RESULT 1)->

This is the comparison between **original vs reconstructed samples at 0th epoch**->



RESULT 2)->

This compares the **original vs reconstructed graph samples after training is completed**

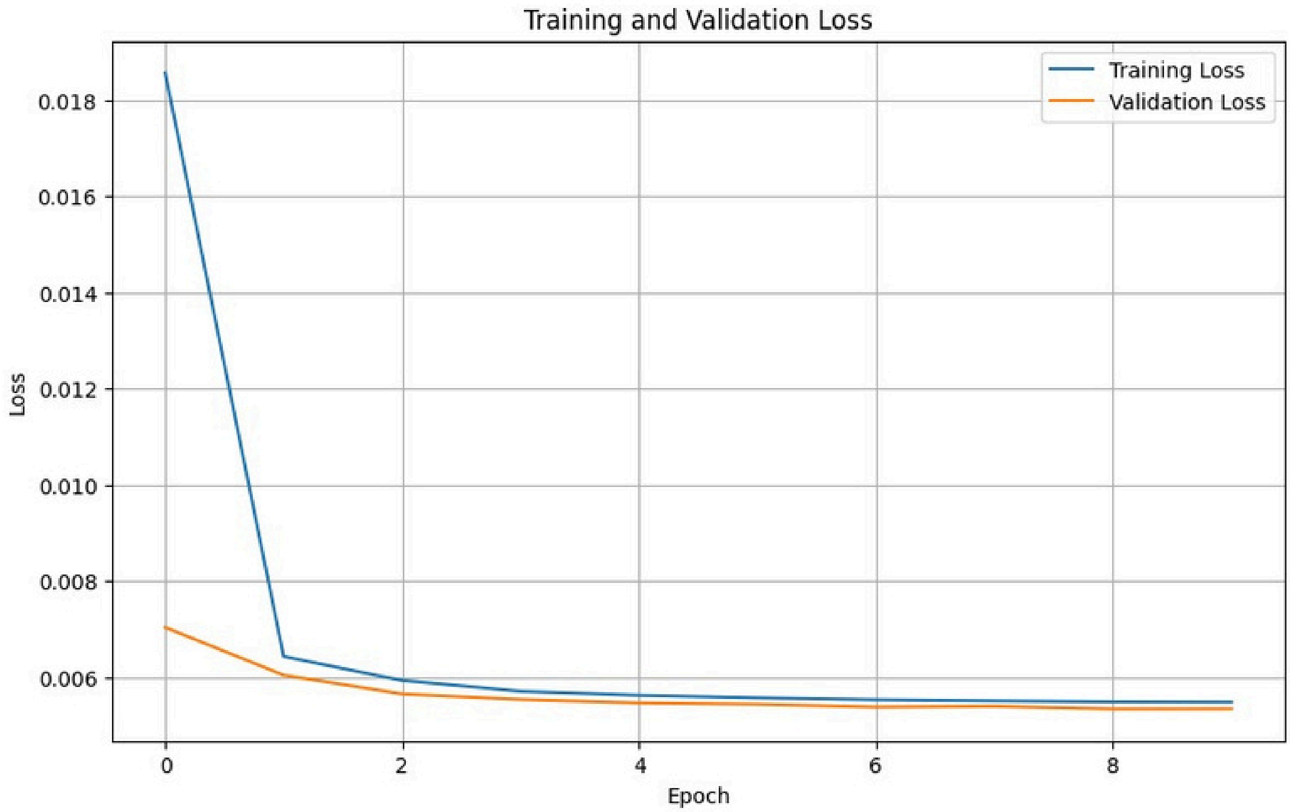
CONCLUSION->

Tables for Organized Comparison

Sample	Initial Epoch (Epoch 0) Observation	Final Epoch (Epoch 10) Observation
Sample 0	Complex network, reconstructed with color gradient, varying accuracy.	Dense cluster, reconstructed similarly, colored nodes aligned.
Sample 1	Dense cluster, reconstructed with gradient, structure preserved.	Elongated shape, reconstructed mirroring original, nodes aligned.
Sample 2	Central cluster, reconstructed with gradient, focusing on density.	Multi-cluster layout, reconstructed with nodes in similar locations.
Sample 3	Horizontal layout, reconstructed with gradient, structure maintained.	Central dense region, reconstructed matching, nodes positioned similarly.
Sample 4	Dense, irregular cluster, reconstructed with pronounced gradient.	Sprawling network, reconstructed maintaining layout, nodes aligned.

The comparison demonstrates that the graph auto-encoder has significantly enhanced its reconstruction capabilities from the initial to the final epoch, effectively learning to simulate detector events with high fidelity. This is promising for applications in fast detector simulation, where accurate and efficient graph representations are essential.

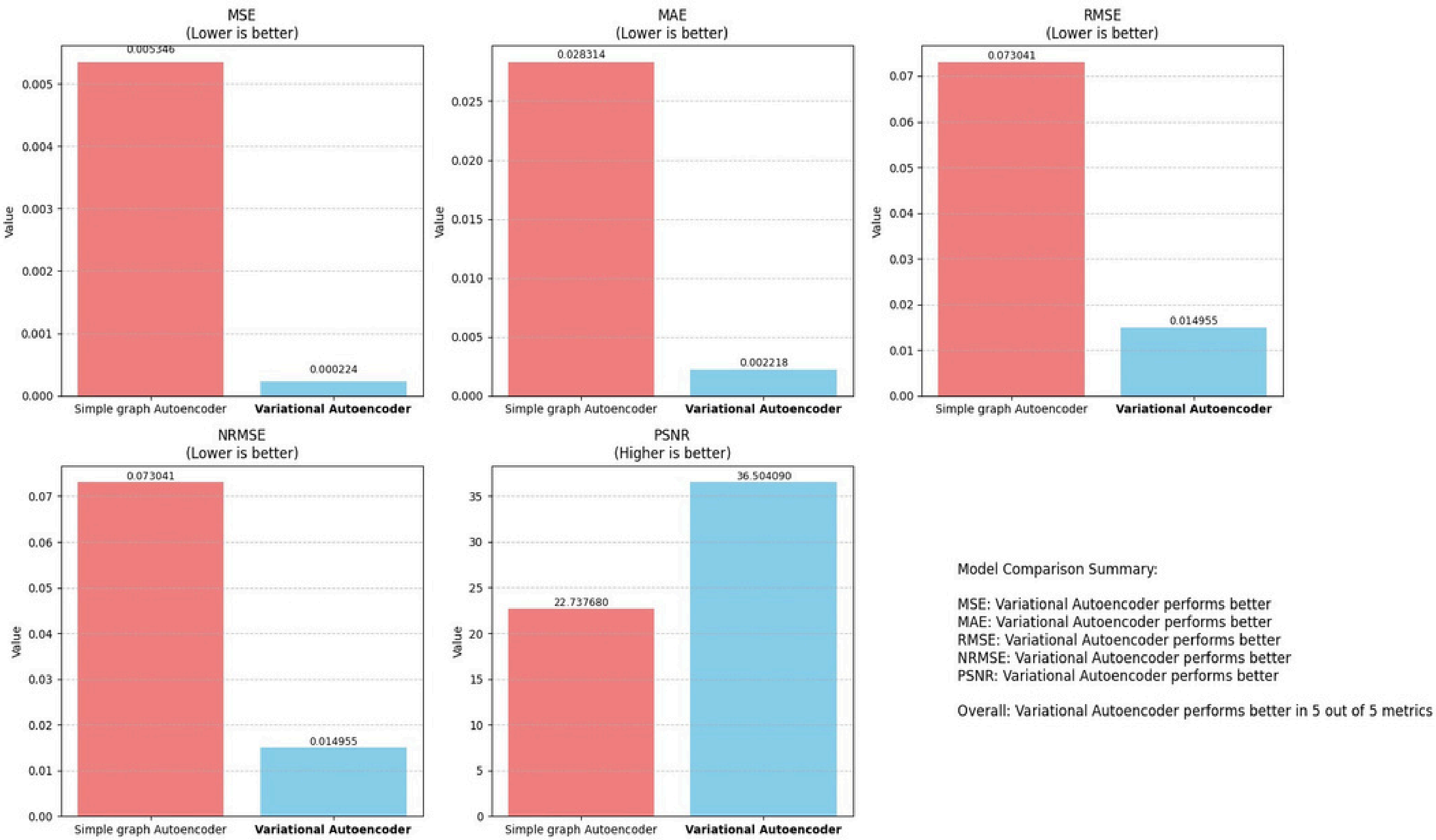
RESULT 3)->
TRAINING/VAL LOSS VS EPOCHS->



Epoch Range	**Training Loss Range**	**Validation Loss Range**	**Observation**
0	0.018	0.007	High initial losses, poor reconstructions
0 to 2	0.018 to 0.006	0.007 to 0.005	Rapid decrease, quick learning
2 to 4	0.006 to 0.005	0.005 to 0.004	Steady decrease, narrowing gap
4 to 8	0.005 to 0.004	0.004 to 0.005	Stable, converged, good generalization

This table summarizes the loss behavior across epochs, highlighting the training process's progression.

COMPARISION WITH VARIATIONAL AUTOENCODER->



Metric	Simple Graph AE	Variational AE	Improvement Factor (VAE vs. Simple)
MSE	0.005346	0.000224	≈ 23.86 (lower)
MAE	0.028314	0.002218	≈ 12.76 (lower)
RMSE	0.073041	0.014955	≈ 4.88 (lower)
PSNR	22.737680	36.504090	≈ 1.6 (higher)
NRMS	0.073041	0.014955	≈ 4.88 (lower)

This table summarizes the metrics and highlights the VAE's superior performance, with lower error metrics and higher PSNR.

CONCLUSION->

The variational auto-encoder outperforms the simple graph auto-encoder across all evaluated metrics, with significant improvements in MSE (23.86 times lower), MAE (12.76 times lower), RMSE and NRMSE (4.88 times lower), and PSNR (1.6 times higher). This suggests the VAE's probabilistic modeling is more suitable for capturing the variability in graph data for fast detector simulation, making it a preferred choice for such applications.