Electron/Photon Classification Task(common task)

Dataset Understanding and Preprocessing

I began by analyzing the provided HDF5 files containing electron and photon data. Each particle sample was represented as a 32x32 image with two channels. After loading the data, I performed exploratory data analysis to understand its characteristics better:

- I visualized sample images from both particle types across both channels
- I analyzed pixel density distributions to identify where signal information was most concentrated
- I computed statistical metrics for both channels to understand data distribution characteristics

This analysis revealed that electrons and photons have distinguishable patterns in their energy deposits, making the classification task feasible using convolutional neural networks.

For preprocessing, I:

1. Combined both datasets and created binary labels (0 for electrons, 1 for photons)
2. Shuffled the data for randomization
3. Split the dataset into training (80%), validation (10%), and test (10%) sets

Model Architecture

I designed a convolutional neural network inspired by ResNet-15 principles but optimized for efficiency while maintaining high discriminative power. My architecture features:

- Four convolutional blocks with increasing feature dimensions (32→64→128→256)
- Strategically decreasing kernel sizes (7→5→3→3) to capture different feature scales
- Batch normalization after each convolution for training stability
- Dropout layers for regularization (0.3 in convolutional layers, 0.5 in fully connected layers)
- A global average pooling layer to reduce parameters
- Two fully connected layers for classification (256→64→1)

This architecture balances model capacity with computational efficiency, resulting in approximately 420K trainable parameters—significantly fewer than a standard ResNet-15 while maintaining excellent classification performance.

Training Methodology

I implemented a robust training methodology with the following components:

- Binary Cross-Entropy with Logits Loss as the optimization objective
- Adam optimizer with weight decay (L2 regularization) to prevent overfitting
- Cosine annealing learning rate scheduler to progressively reduce the learning rate
- Model checkpointing to save the best model based on validation accuracy
- Comprehensive performance tracking (loss, accuracy, learning rate)

Training was conducted for 50 epochs with a batch size of 2048, striking a balance between training speed and convergence stability. I implemented early stopping based on validation accuracy to prevent overfitting.

Results and Evaluation

The trained model achieved the following performance metrics on the test set:

- Accuracy: 72.55%
- Precision: 74.13%
- Recall (Sensitivity): 69.57%
- Specificity: 75.54%
- F1 Score: 71.78%

The best validation accuracy achieved was 72.50%, indicating that the model generalizes well to unseen data with minimal overfitting(when I tried with training (80%), validation (10%), and test (10%) sets).

Analysis and Insights

I conducted additional analyses to gain deeper insights into model performance:

1. Feature map visualization to understand what patterns the model was detecting
2. Performance analysis by image intensity, reveals that classification accuracy varies based on signal strength
3. Sample prediction visualization to identify challenging cases

These analyses highlight that the model performs better on particles with clearer energy deposit patterns and struggles more with low-energy deposits or particles with ambiguous signatures.

Challenges and Solutions

During implementation, I encountered several challenges:

1. **Data imbalance**: I addressed this by ensuring equal representation of both classes in the dataset.
2. **Feature extraction efficiency**: I optimized the architecture to balance computational cost with discriminative power.
3. **Overfitting**: I incorporated dropout, weight decay, and learning rate scheduling to regularize the model effectively.

Conclusion

The implemented classifier achieves a solid accuracy of 72.55% on the test set, providing reliable discrimination between electrons and photons based solely on their detector signatures. The model architecture balances efficiency and performance, making it suitable for deployment in high-energy physics analysis pipelines.

Future improvements could include:

- Implementing data augmentation strategies specific to particle physics
- Exploring attention mechanisms to focus on discriminative regions
- Fine-tuning hyperparameters using Bayesian optimization