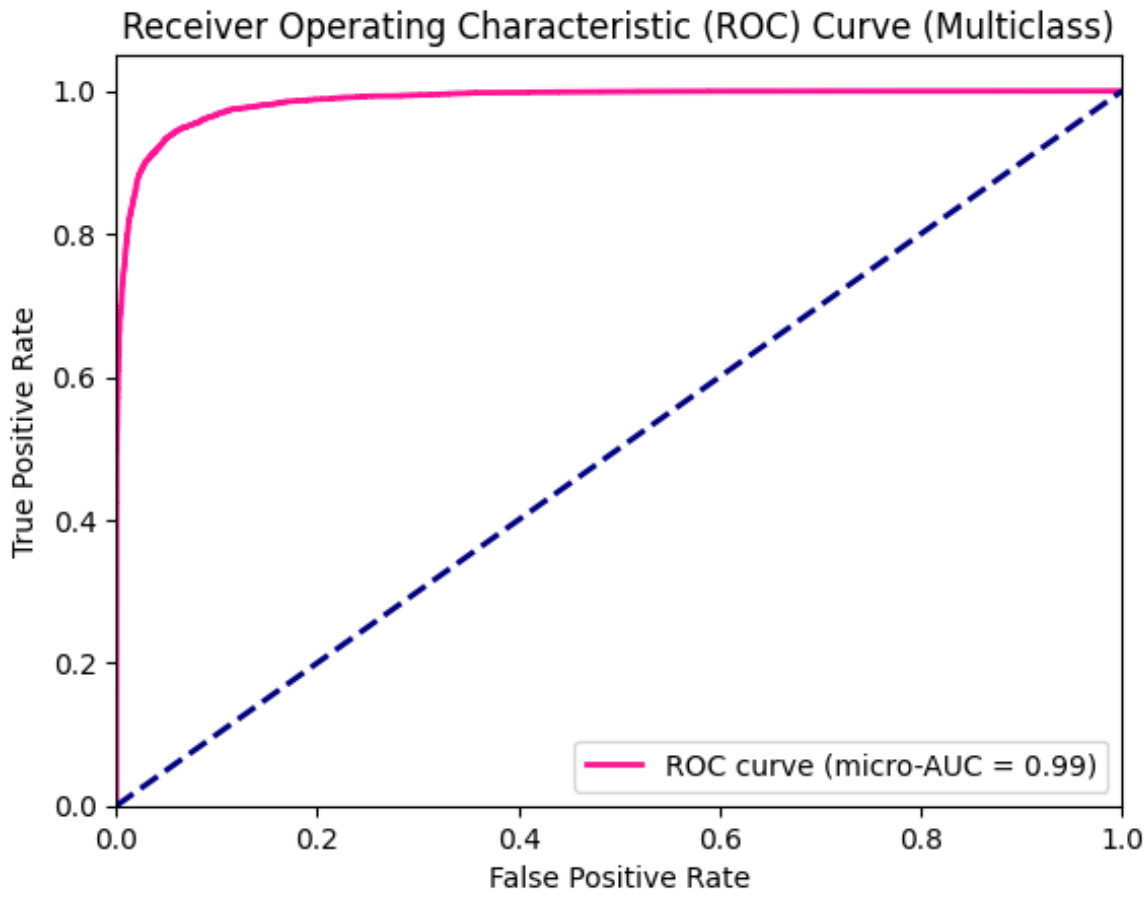
Common Test I. Multi-Class Classification

Task: Build a model for classifying the images into lenses using PyTorch or Keras. Pick the most appropriate approach and discuss your strategy.

Result(notebook):

Trained the ResNet-18 on the given dataset with a 5-fold cross-validation with each 5 epochs on 3*10k training images and used 3*2.5k for test and reached the ROC-AUC score of 0.99 on test data.



I used ResNet-18 as it can capture more intricate patterns and features in the data, leading to better generalization and higher accuracy. And the model should not be much deep as it will overfit for grayscale images. The ability of ResNets to carry many useful features make ResNet-18 most appropriate choice. Also from various research papers ResNet-18 make the best results for lens classification.

Specific Test VI. SSL on Real Dataset:

Task: Build a Self-Supervised Learning model for classifying the images into lenses using PyTorch or Keras. Pick the most appropriate approach and discuss your strategy.

Approaches:

Use the 215 images provided to train the self-supervised learning models in two different ways with 5-fold cross-validation set in pytorch. Contrastive learning with NT-Xent loss function used in both approaches.

1. **ResNet-18 with Rotation pretext training(notebook):**
   In this model ResNet-18 is used as the backbone architecture for feature extraction and a projection head is applied after this. Rotational pretext learning is used for training the model.
2. **CNN+Self attention with Gaussian Noise pretext training(notebook):**
   Only a small CNN architecture is used as a backbone with self-attention and a projection head after this. Self-attention will help the model to focus more on the galaxy in the large background. Gaussian Noise is used for pretext learning in this case.

Fine – Tuning:

Both models were then fine-tuned on the same image dataset of 215 images with 5-fold cross-validation. Backbone architecture weights were frozen replacing the projection head with the classification head in the same above notebooks. It was used for classification into two classes lens and not lens.

Results:

Then ROC-AUC metrics were evaluated on the same 215 image dataset. The results were acceptable for the given very small dataset.

2nd approach gives low score as the backbone architecture was way very small compared to the ResNet-18 used in 1st approach.

The second approach opens further possibilities for combining the different CNNs and self–attention in different ways.

The backbone architectures can be also replaced with VITs and hybrid architectures. Also a combination of different types of pretext training in a large dataset would help to learn the representation of the images more effectively.