Deciphering the Chest : Deep Learning for Pathological Insights with Pre-trained Models.

Summary

This project aims to develop a deep learning system for multi-label classification of thoracic pathologies in chest X-ray images using the ChestMNIST dataset. It seeks to identify effective architectures and techniques for accurate diagnosis, addressing the need for automated tools to assist healthcare professionals and potentially enhancing diagnostic accuracy and patient outcomes. This endeavor is crucial for improving healthcare delivery by providing timely and accurate diagnoses, vital for effective patient management and treatment planning.

Objectives

To classify ChestMNIST dataset into multiple thoracic pathologies based on X-ray images.

To investigate the performance of various deep learning architectures, including DenseNet, GoogleNet, ResNeXt, VGG16, ResNet-18, and ResNet-50, in accurately identifying thoracic pathologies from X-ray images.

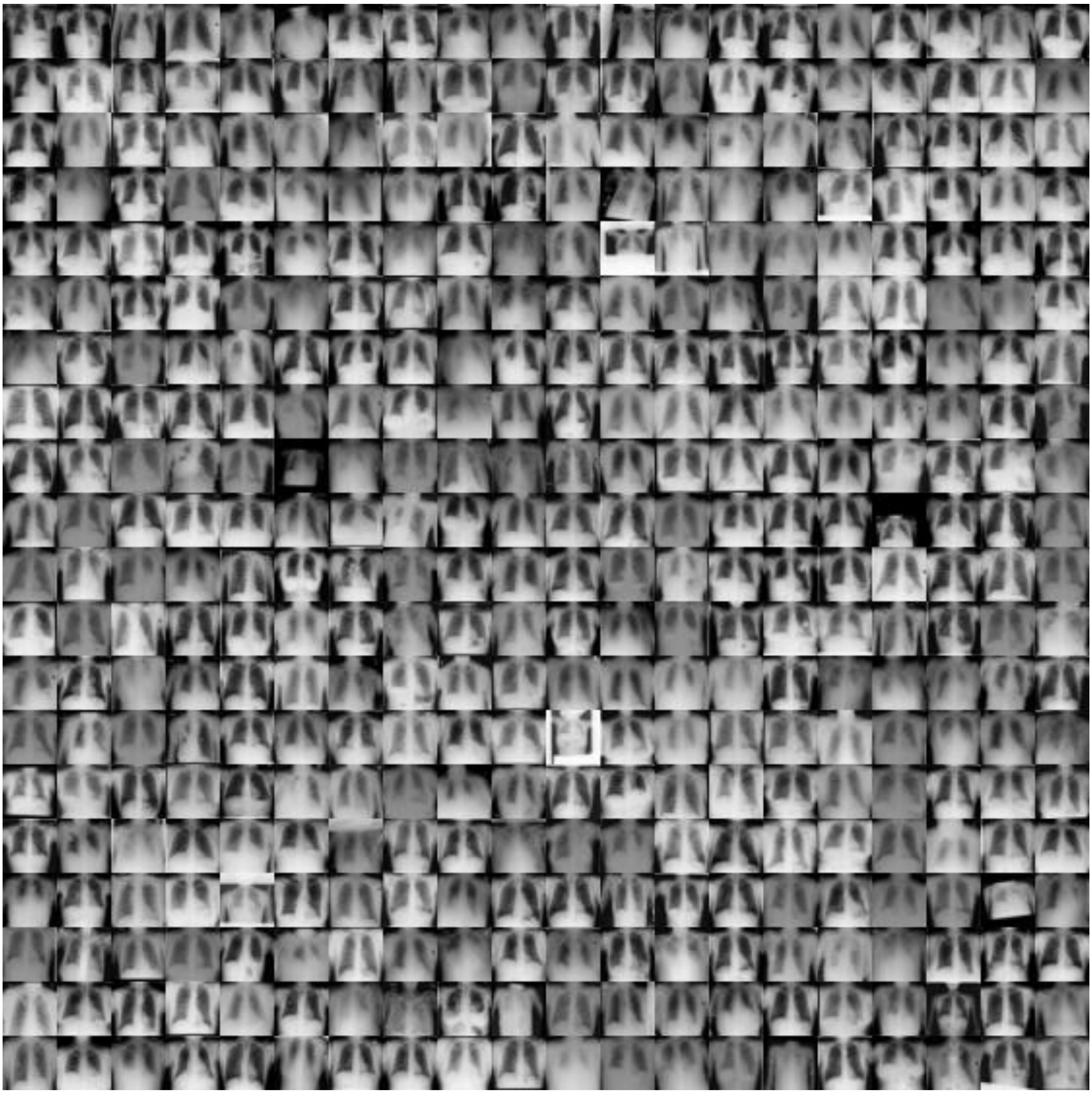
About the dataset

ChestMNIST dataset is a collection of chest X-ray images that have been converted into a format similar to the MNIST dataset, making it ideal for image classification tasks.

Each image in the dataset represents one of the following classes: atelectasis, cardiomegaly, effusion, infiltration, mass, nodule, pneumonia, pneumothorax, consolidation, edema, emphysema, fibrosis, pleural thickening, hernia.

This dataset serves as a valuable resource for developing and testing deep learning models for chest X-ray analysis.

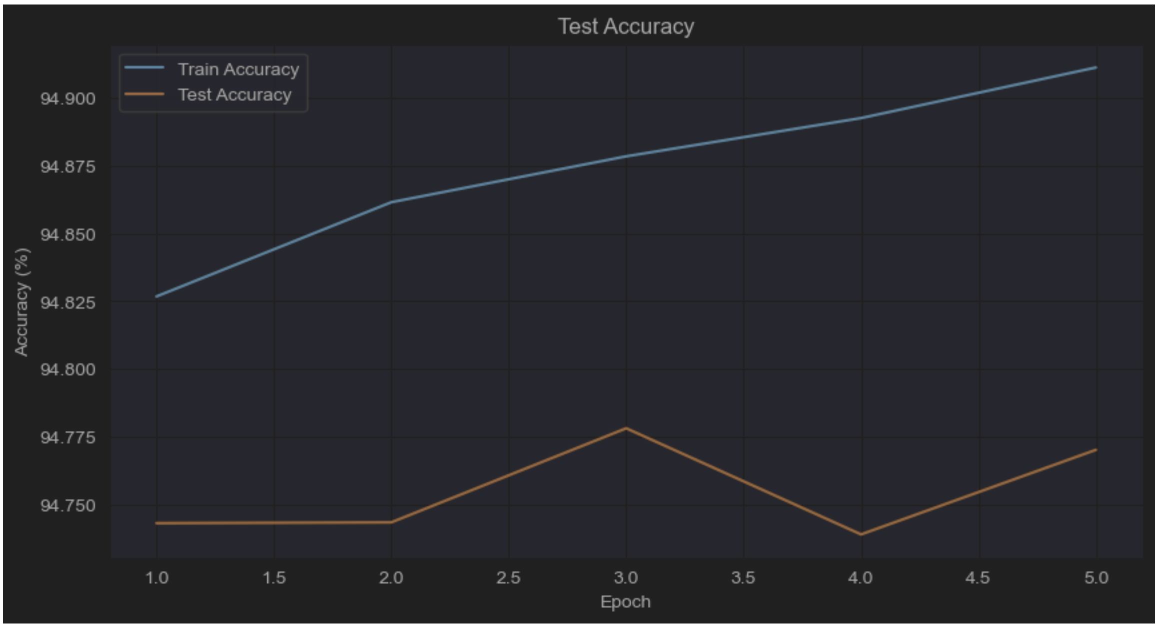
The images in chest dataset are displayed and visualized to understand the various images.

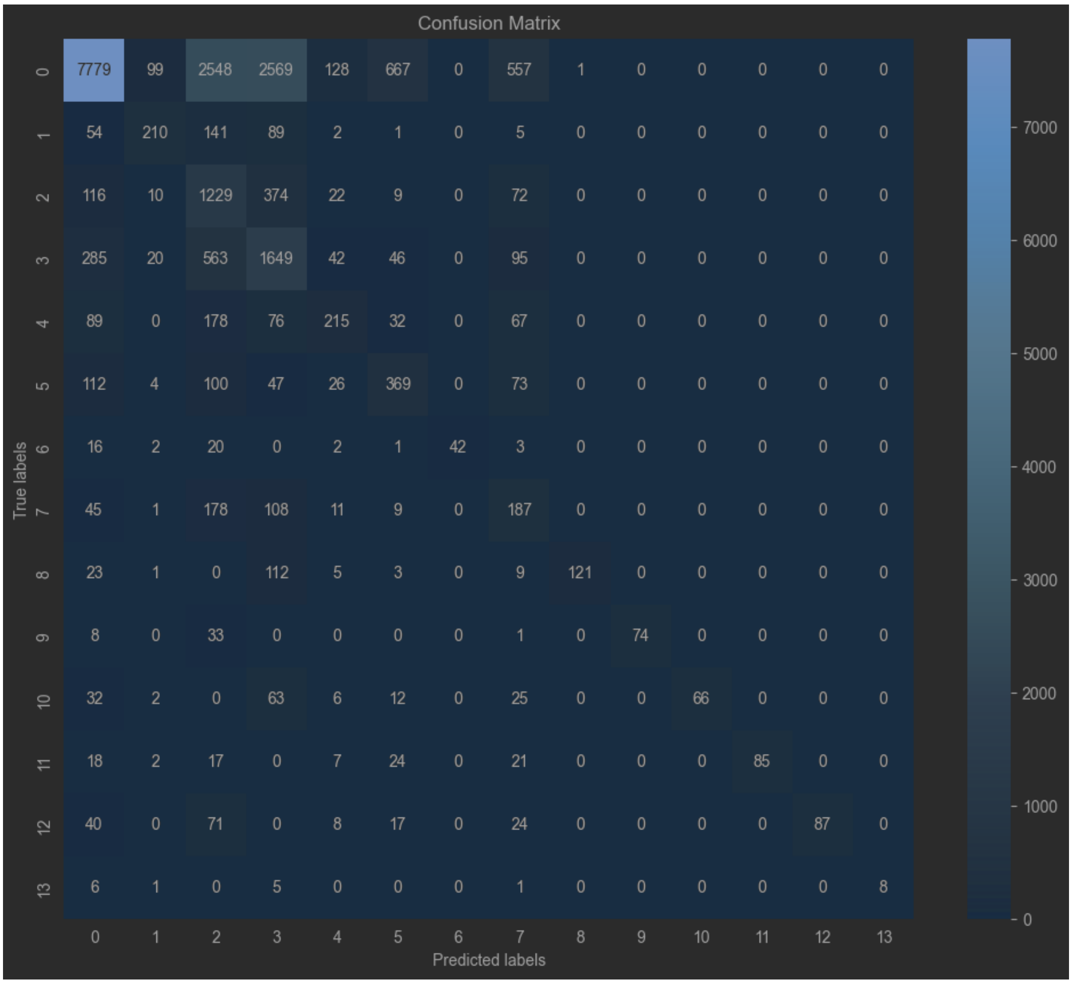


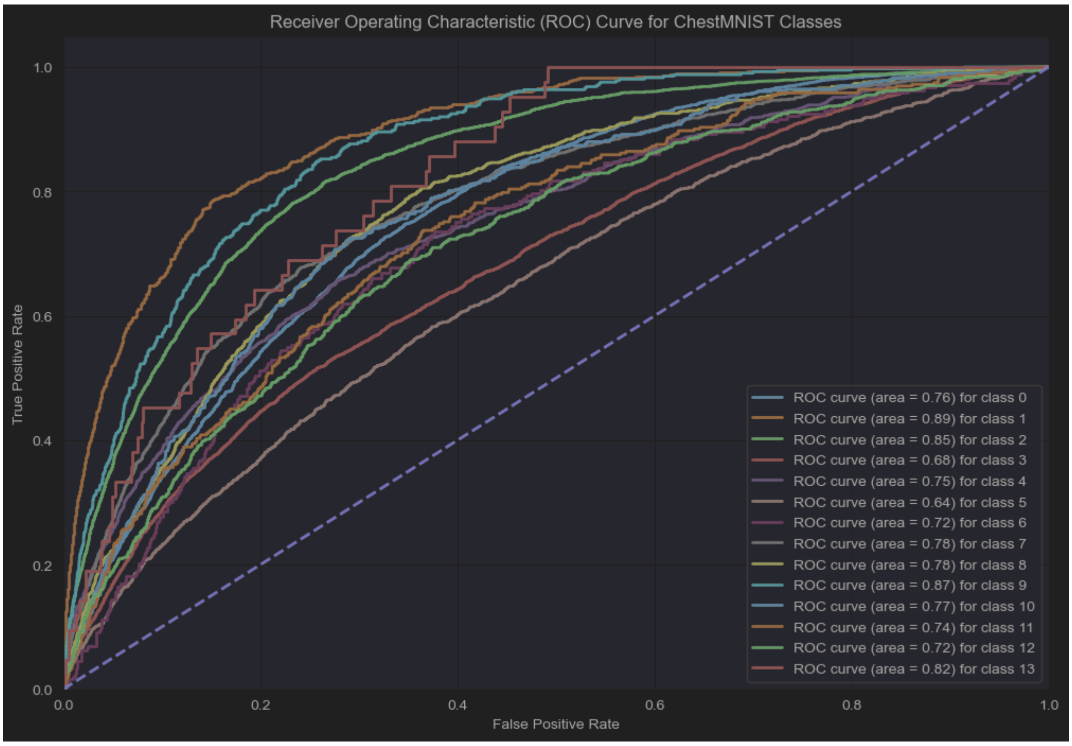
Pre-trained Models

ResNet18 :

- Architecture: Residual Network with 18 layers
- Key Strengths: Effective residual connections, robust feature extraction, good performance on various tasks
- Potential Drawbacks: Relatively shallow architecture compared to deeper variants
- Train Loss: 0.1582, Train Acc: 94.91%, Test Loss: 0.1704, Test Acc: 94.77%



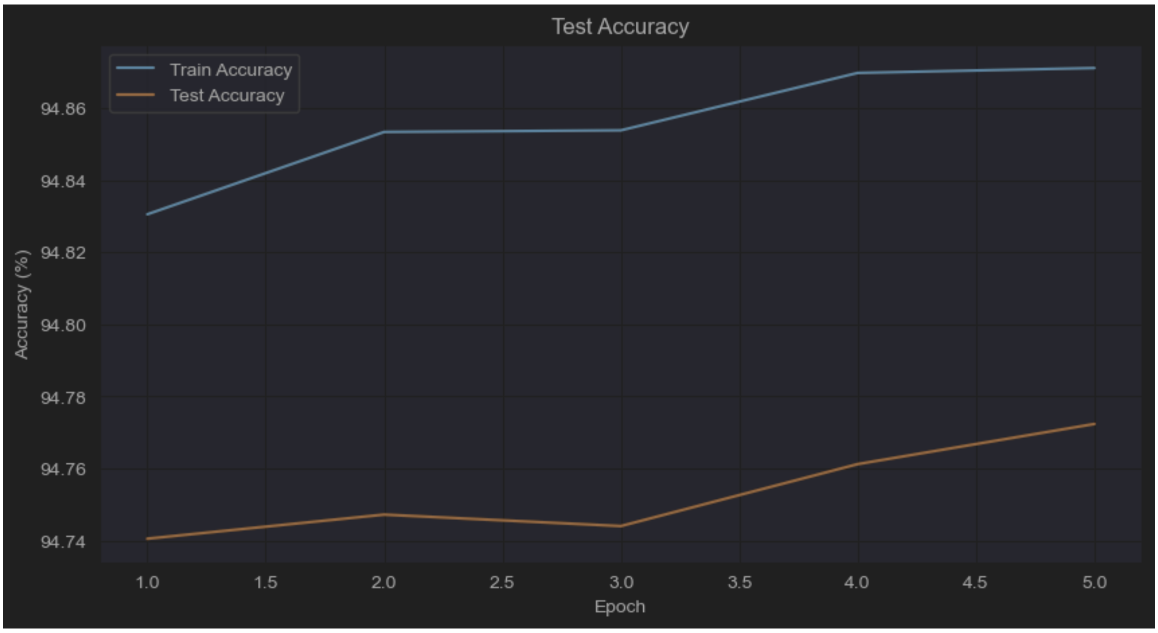


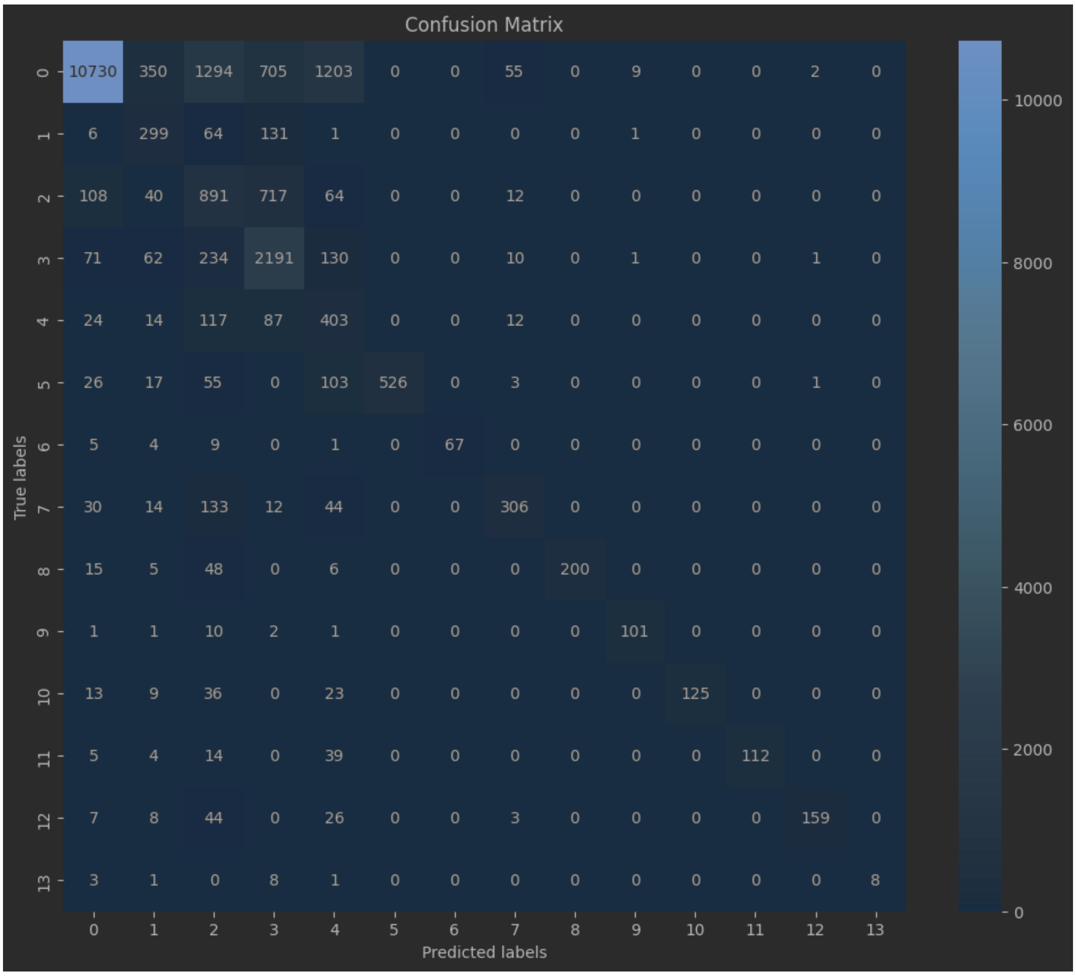


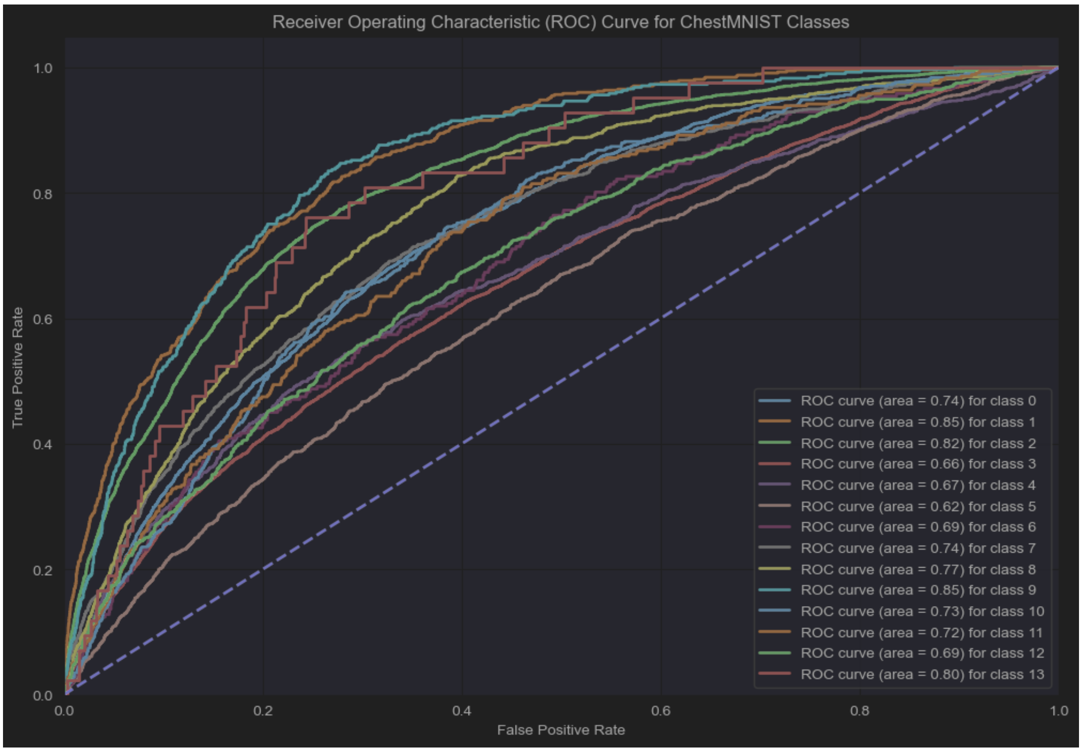
We are able to visualize the classification of data under each class and the extent of correctness using the area under each class of Receiver Operating Curve

ResNet50 :

- Architecture: Residual Network with 50 layers
- Key Strengths: Deeper architecture with increased representational capacity, proven performance on various tasks
- Potential Drawbacks: Higher computational complexity, potential for overfitting on smaller datasets
- Train Loss: 0.1658, Train Acc: 94.87%, Test Loss: 0.2229, Test Acc: 94.77%



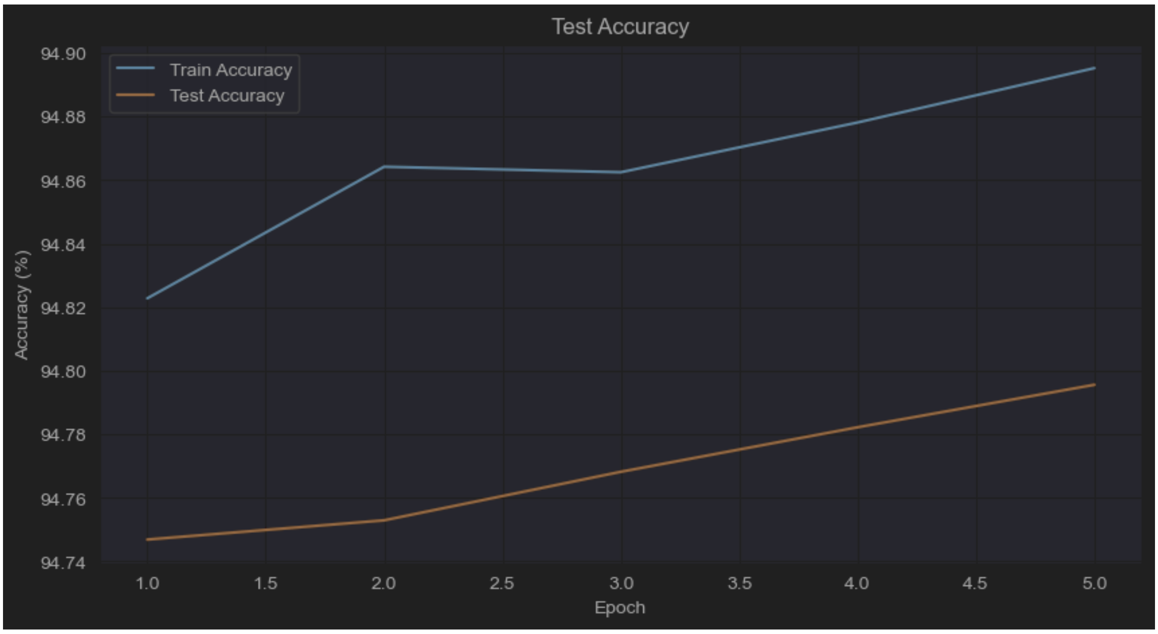




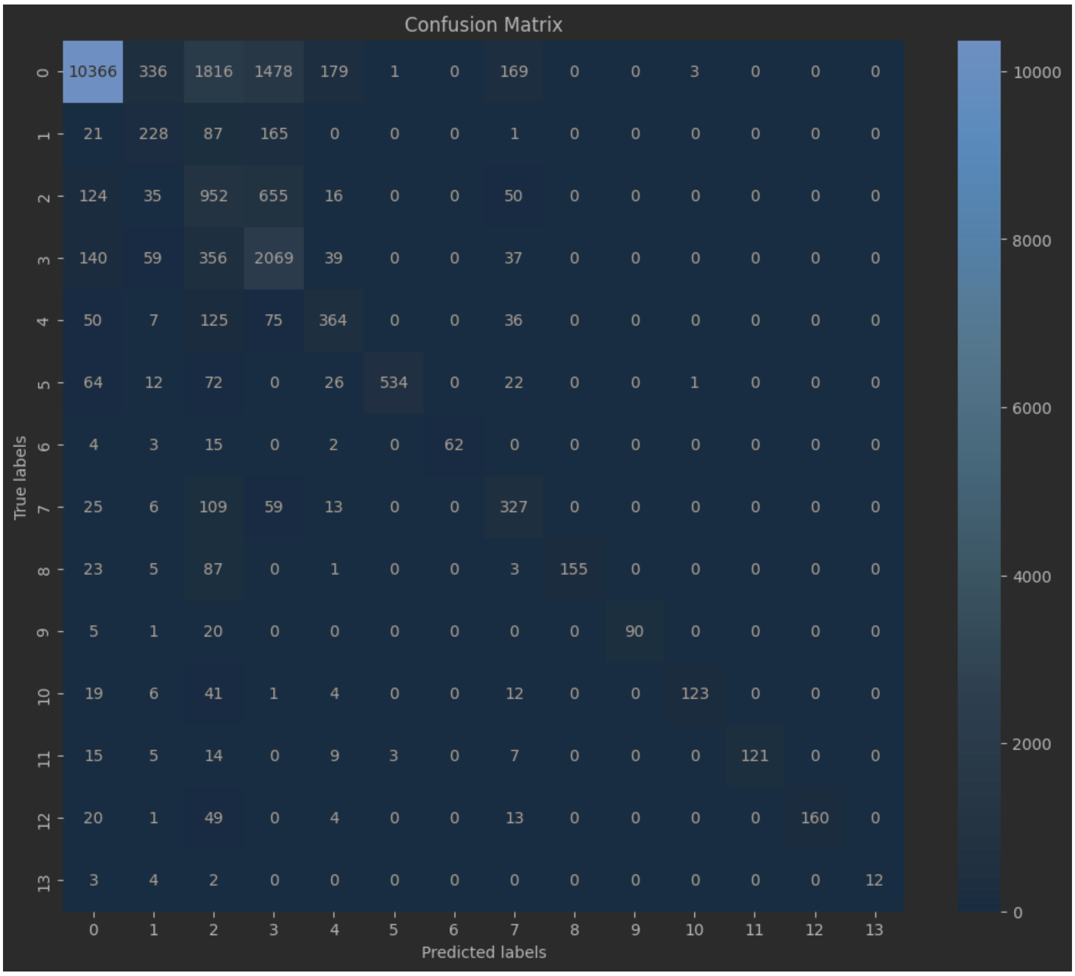
We are able to visualize the classification of data under each class and the extent of correctness using the area under each class of Receiver Operating Curve

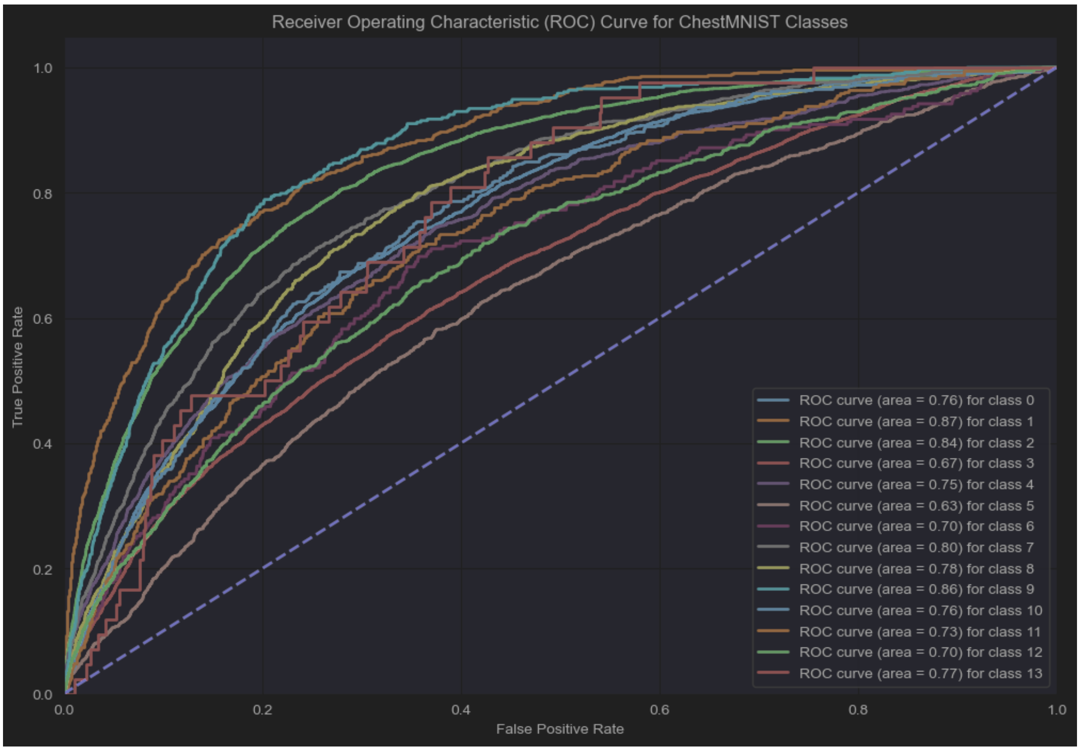
DenseNet121 :

- Architecture: Dense Convolutional Network with 121 layers
- Key Strengths: Feature reuse through dense connections, efficient parameter utilization, alleviates vanishing gradient problem
- Potential Drawbacks: Memory-intensive due to dense connections, potential for overfitting on smaller datasets
- Train Loss: 0.1613, Train Acc: 94.90%, Test Loss: 0.2602, Test Acc: 94.80%



We could able to infer that loss decreases with increase in episodes, and accuracy increase with increase in number of episodes

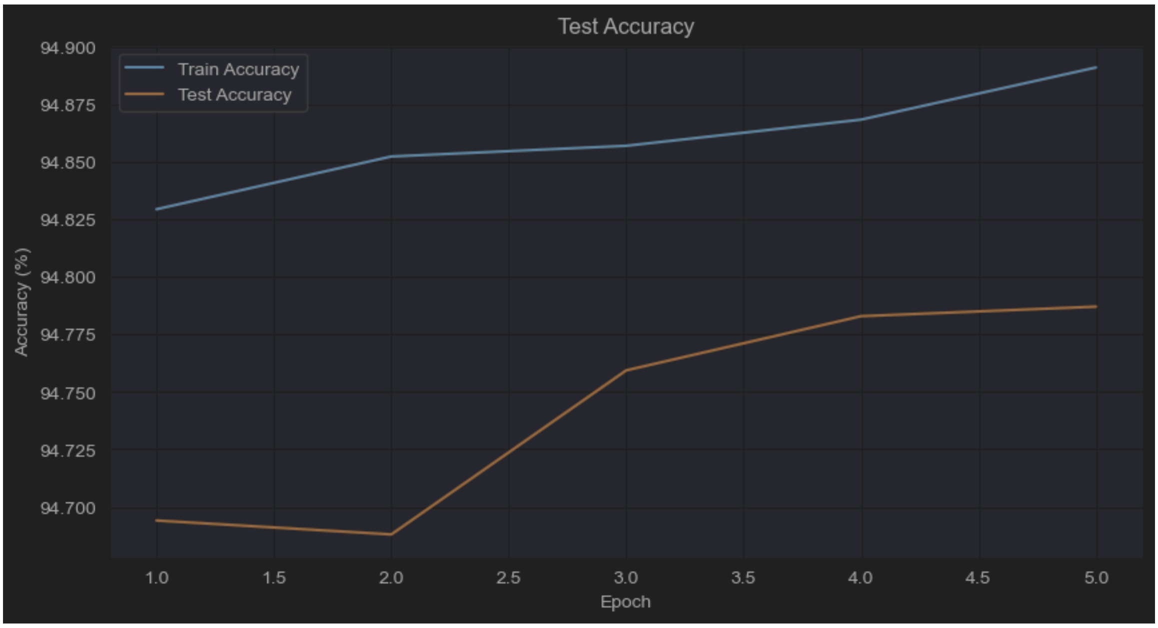




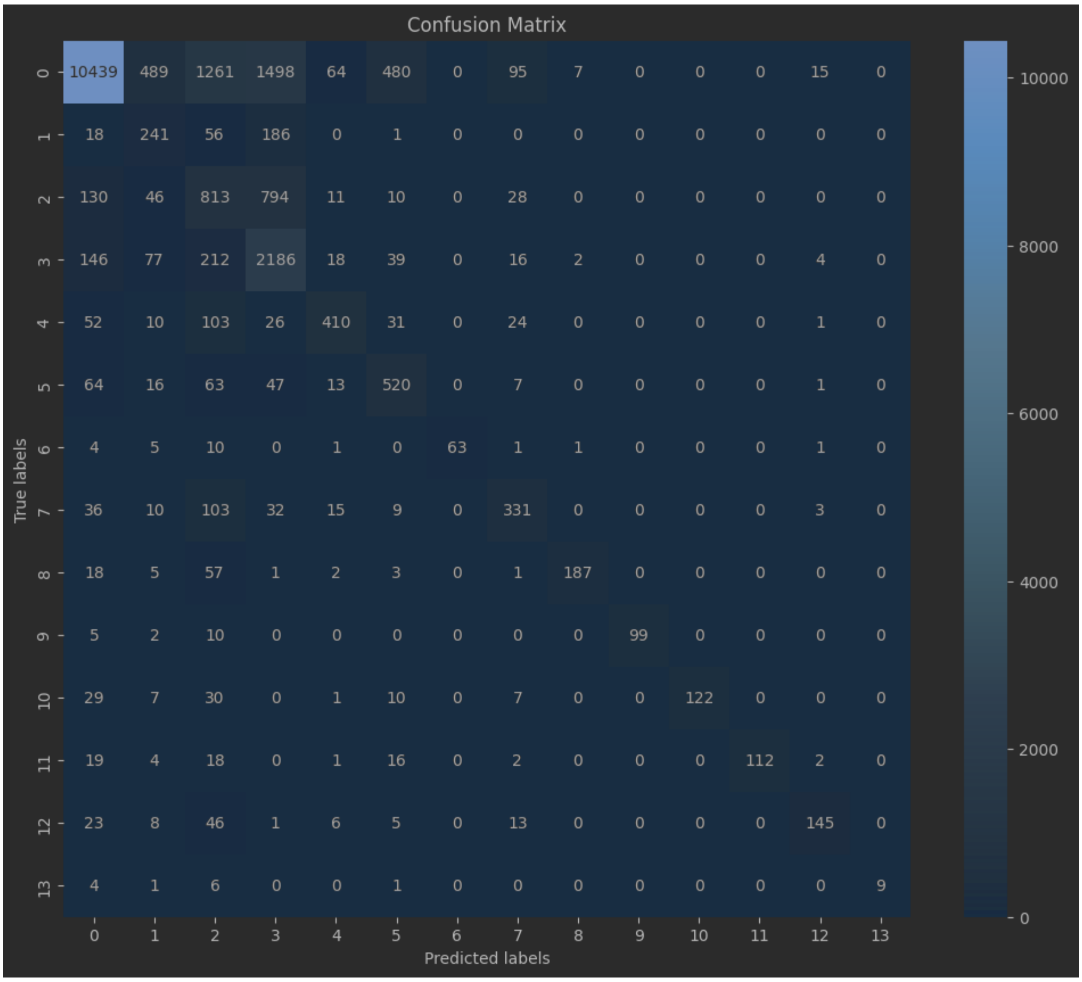
We are able to visualize the classification of data under each class and the extent of correctness using the area under each class of Receiver Operating Curve

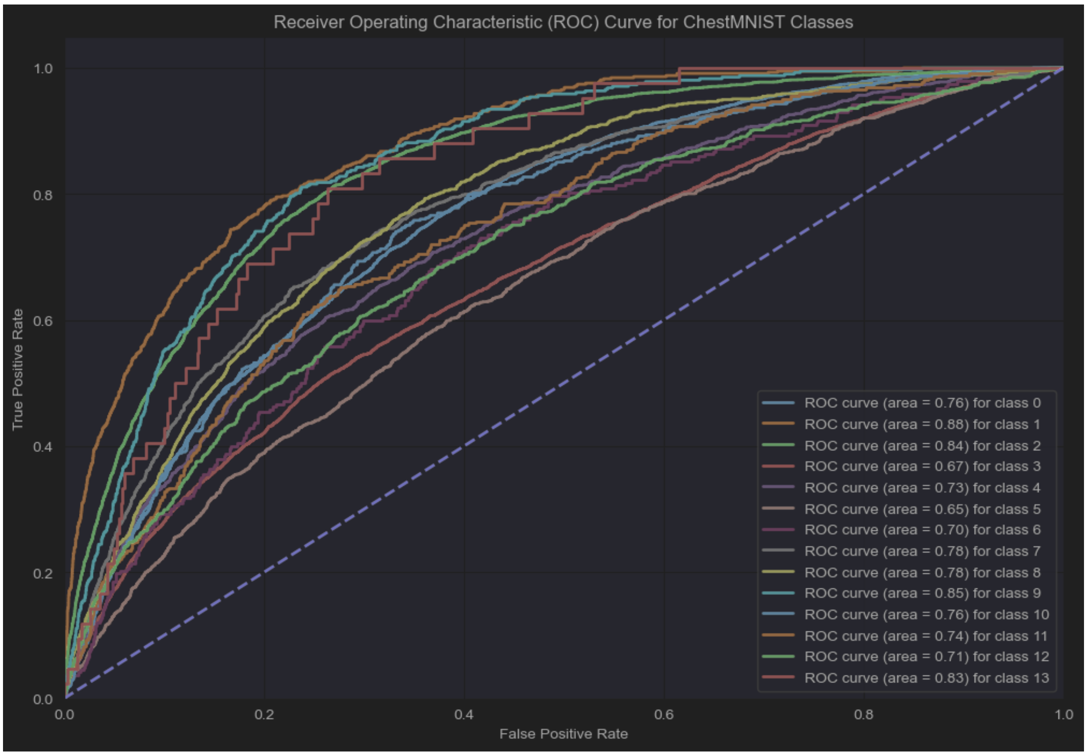
ResNeXt50 :

- Architecture: Aggregated Residual Transformations with 50 layers
- Key Strengths: Improved representational capacity through cardinality, better generalization performance
- Potential Drawbacks: Increased computational complexity, potential for overfitting on smaller datasets
- Train Loss: 0.1628, Train Acc: 94.89%, Test Loss: 0.1949, Test Acc: 94.79%



We could able to infer that loss decreases with increase in episodes, and accuracy increase with increase in number of episodes

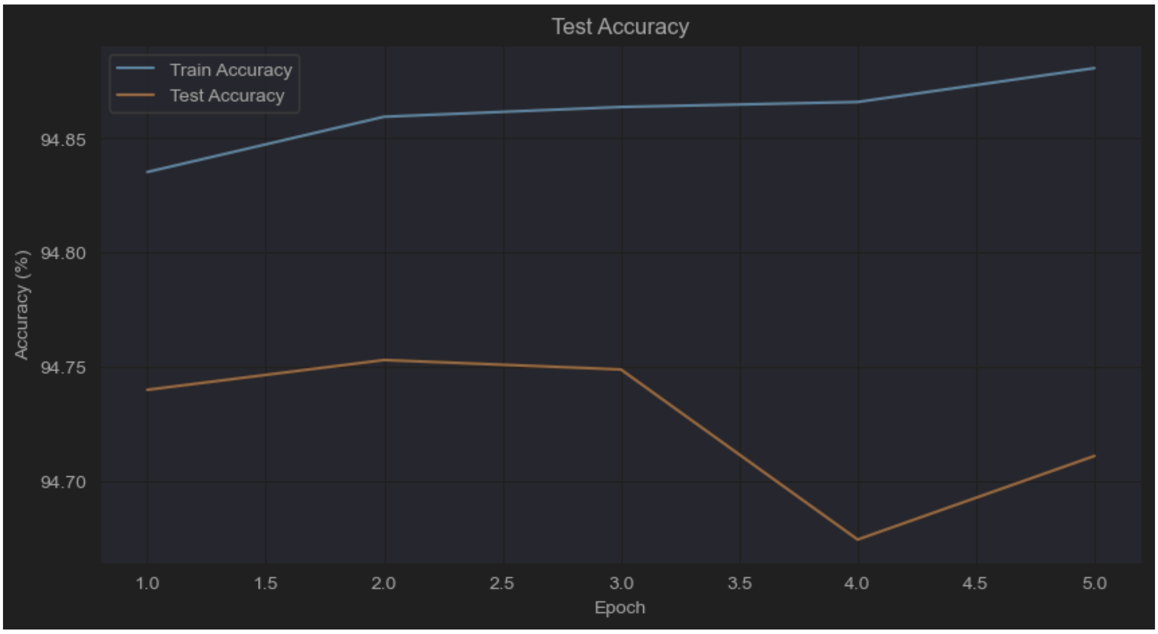




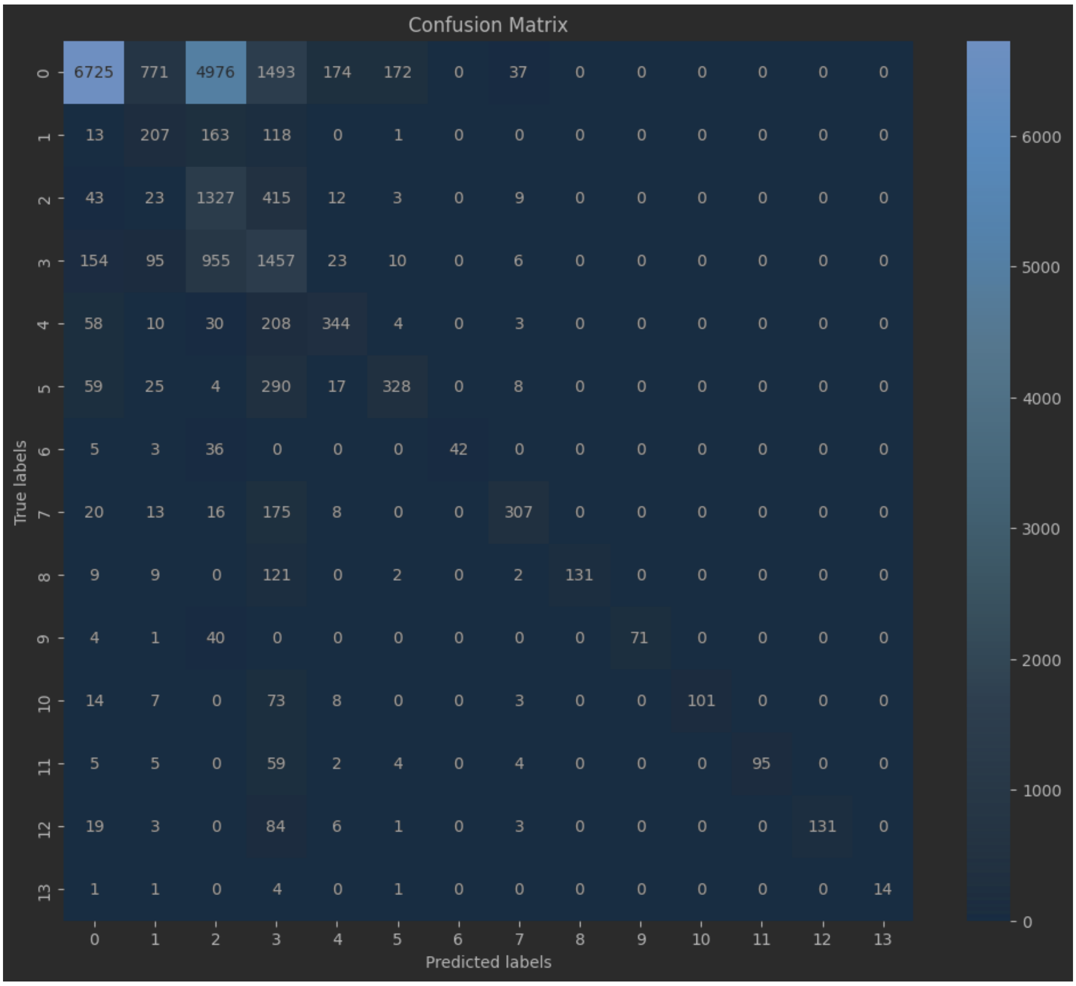
We are able to visualize the classification of data under each class and the extent of correctness using the area under each class of Receiver Operating Curve

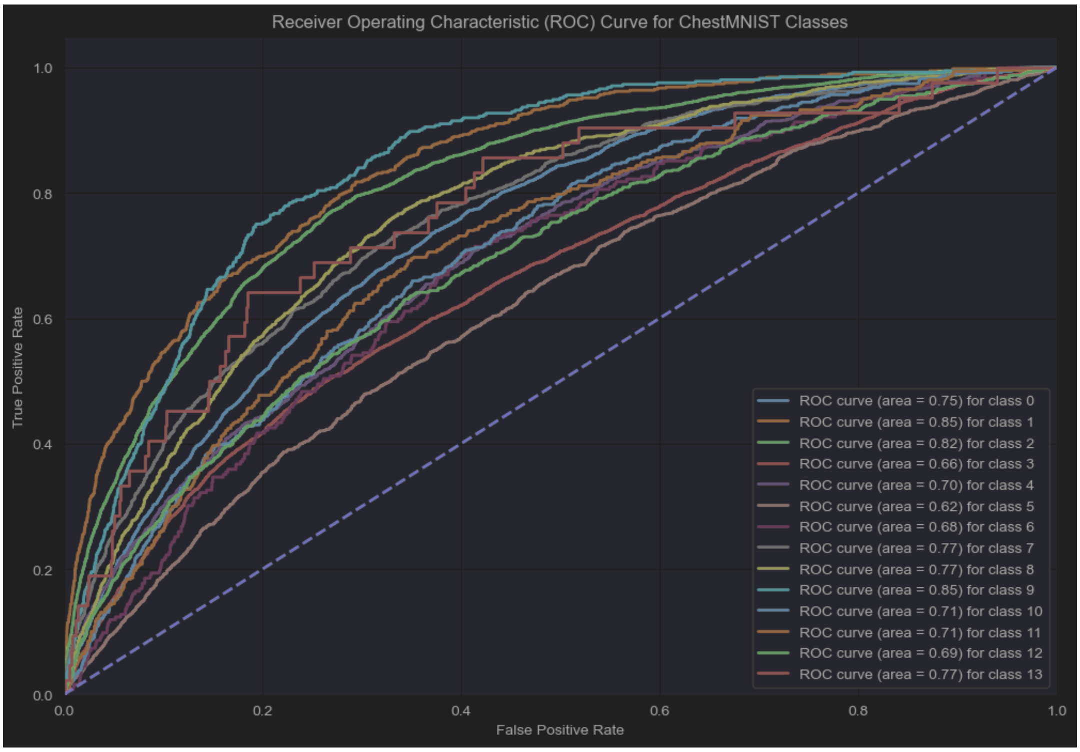
GoogLeNet :

- Architecture: Inception Network with auxiliary classifiers
- Key Strengths: Efficient use of computational resources, improved performance through auxiliary classifiers
- Potential Drawbacks: Complex architecture, potential for overfitting on smaller datasets
- Train Loss: 0.1648, Train Acc: 94.88%, Test Loss: 0.1902, Test Acc: 94.71%



We could able to infer that loss decreases with increase in episodes, and accuracy increase with increase in number of episodes



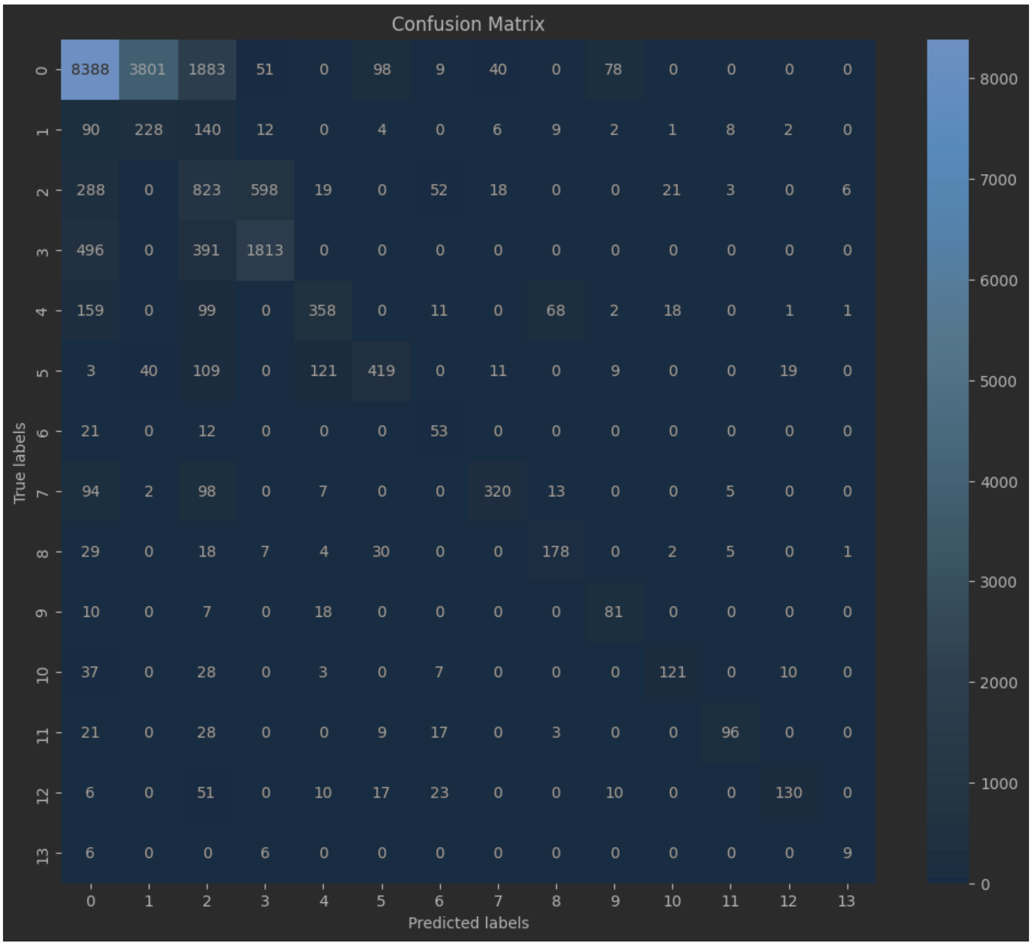


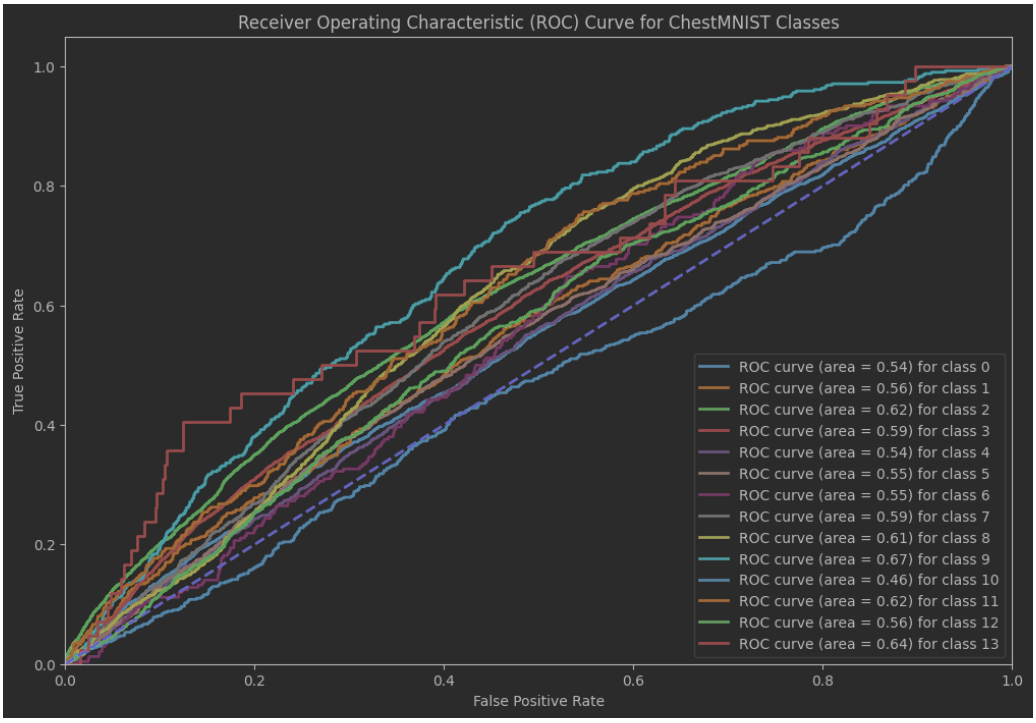
We are able to visualize the classification of data under each class and the extent of correctness using the area under each class of Receiver Operating Curve

VGG16 :

- Architecture: Very Deep Convolutional Network with 16 layers
- Key Strengths: Simple and straightforward architecture, effective for feature extraction
- Potential Drawbacks: Computationally expensive, potential for overfitting on smaller datasets
- Train Loss: 0.1835, Train Acc: 94.86%, Test Loss: 0.2765, Test Acc: 94.74%

We could able to infer that loss decreases with increase in episodes, and accuracy increase with increase in number of episodes





We are able to visualize the classification of data under each class and the extent of correctness using the area under each class of Receiver Operating Curve

Results:

Experimental Setup

CPU : R9 5900X

GPU : 3080 Ti

RAM : 32 GB 3.6 GHz

Pytorch version : 2.2.2

CUDA version : 12.1

IDE : DataSpell 2024.1

Comparisson:

Model	Epoch	Train Loss	Accuracy (%)	Test Loss	Accuracy (%)	Precision	Recall	F-1 Score
ResNet18	5	0.1582	94.91	0.1704	94.77	0.9231	0.9474	0.9219
ResNet50	5	0.1658	94.87	0.2229	94.76	0.9298	0.9477	0.9237
DenseNet121	5	0.1613	94.90	0.2602	94.80	0.9293	0.9480	0.9264
ResNeXt50	5	0.1628	94.89	0.1949	94.79	0.9326	0.9479	0.9239
GoogLeNet	5	0.1648	94.88	0.1902	94.71	0.9255	0.9471	0.9274
VGG16	5	0.1835	94.86	0.2765	94.74	0.8976	0.9474	0.9219

Observations from the Result:

1. Overall Performance: All models achieve high accuracy (>94%) on both the training and test sets, indicating that they have learned the features of the dataset well. This suggests that the models are not overfitting, as the test accuracy is comparable to the training accuracy.

2. Loss Values: The training and test losses are relatively low for all models, which indicates that the models are effectively minimizing the error during training and generalizing well to the test set.

3. Accuracy: DenseNet121, ResNeXt50 and ResNet 18 have the highest accuracies. This suggests that these models are well-suited for the ChestMNIST dataset.

4. Precision, Recall, F1-Score: All models have high precision, recall, and F1-score values.

ResNet50 and GoogLeNet appear to be the most promising models for further optimization and deployment mainly because of their balanced performance across various metrics.

Future Works:

Our future scope would be leveraging attention mechanisms for this Chestminist dataset. Attention mechanisms enable models to focus on relevant regions of an image, allowing for more precise and interpretable predictions. That will be the extension of our Computer Vision project in the next semester.

References:

https://scikit-learn.org/stable/

https://arxiv.org/pdf/1512.03385.pdf - Deep Residual Learning for Image Recognition

https://arxiv.org/pdf/1409.4842.pdf - Going deeper with convolutions.