

Indian Institute of Technology Jodhpur



## ESC-10 Classification using 1D Convolution and Transformer

---

Submitted by

MITESH KUMAR (M23MAC004)

Colab Link:

 M23MAC004.ipynb

- Model Trainer class is responsible for the training of all 4 Architectures
- Calculate loss and accuracy on all train, test and validation datasets print it and log these on (WandB).
- It also plots the train-validation loss and accuracy graph.
- For the test set, it creates the confusion matrix and AUC-ROC Curve.
- Training and validation is done using 4 Fold train-validation.

## Architecture 1

### 1D CNN

It is a 5-layer 1D convolution followed by 1 1-layer fully connected layer.

Activation: ReLU, Sigmoid.

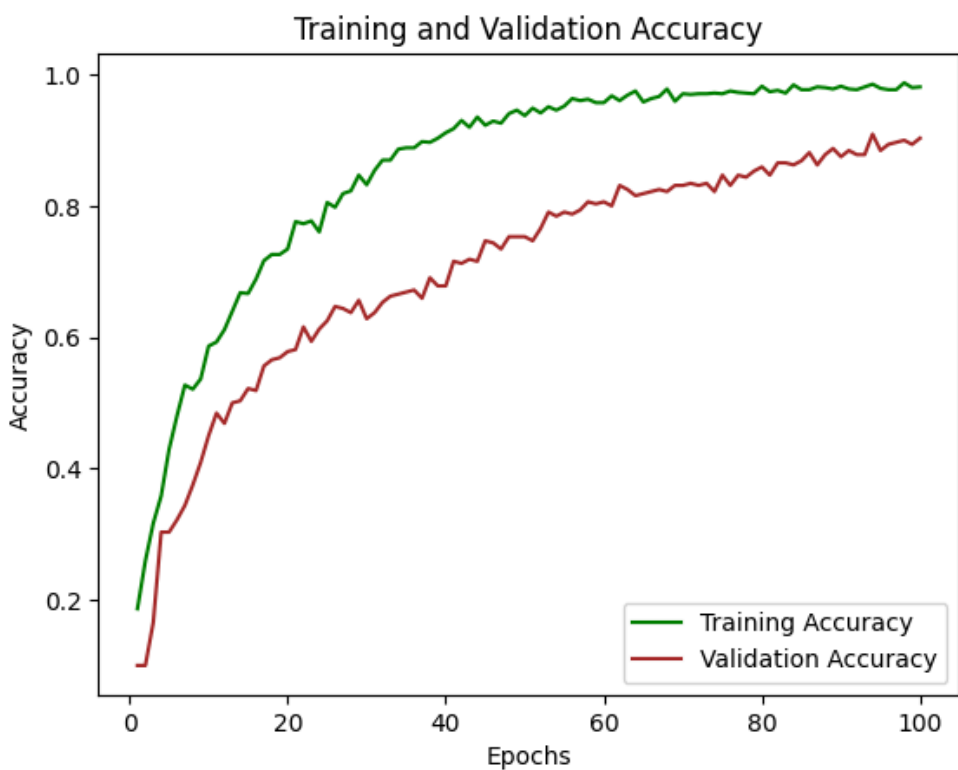
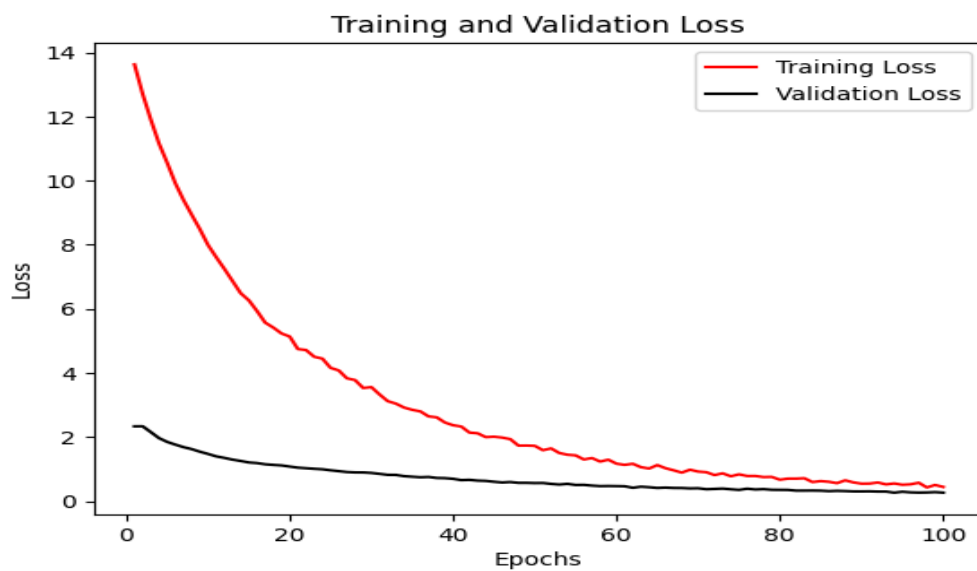
Batch Normalization, Max-pooling, and Dropout are applied.

Learning rate = 0.001, batch size = 40

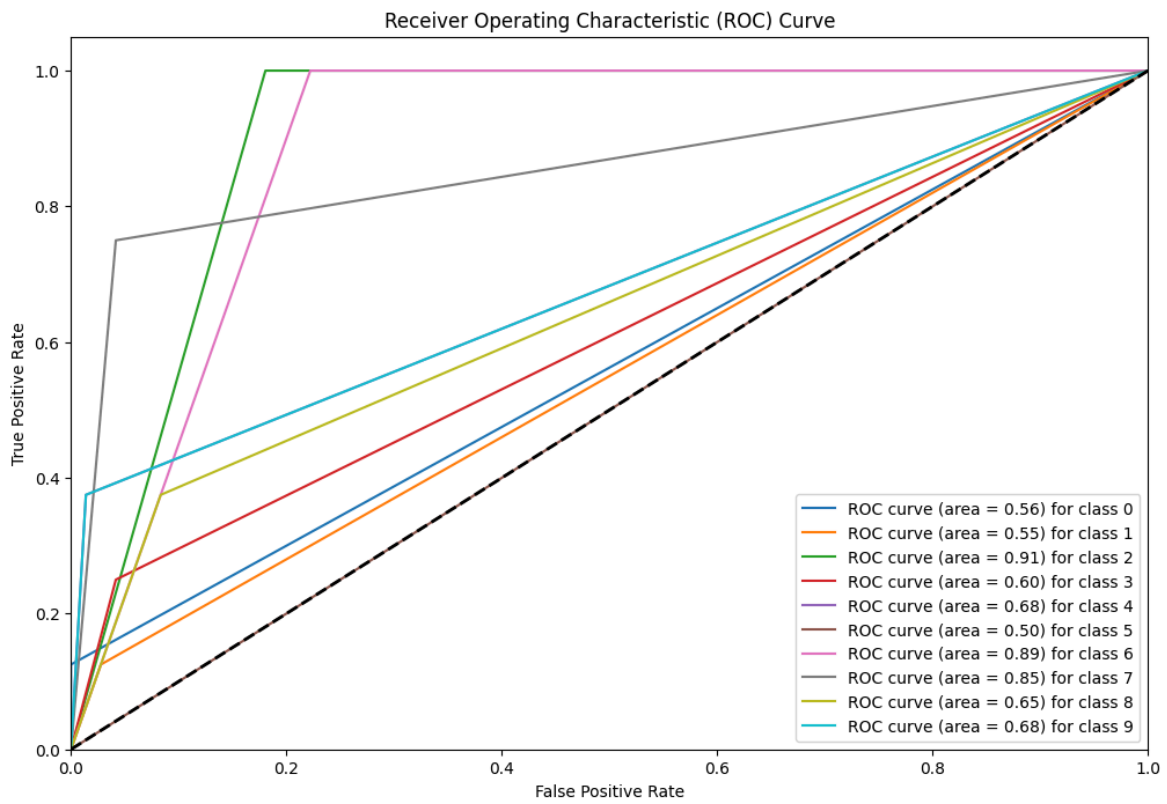
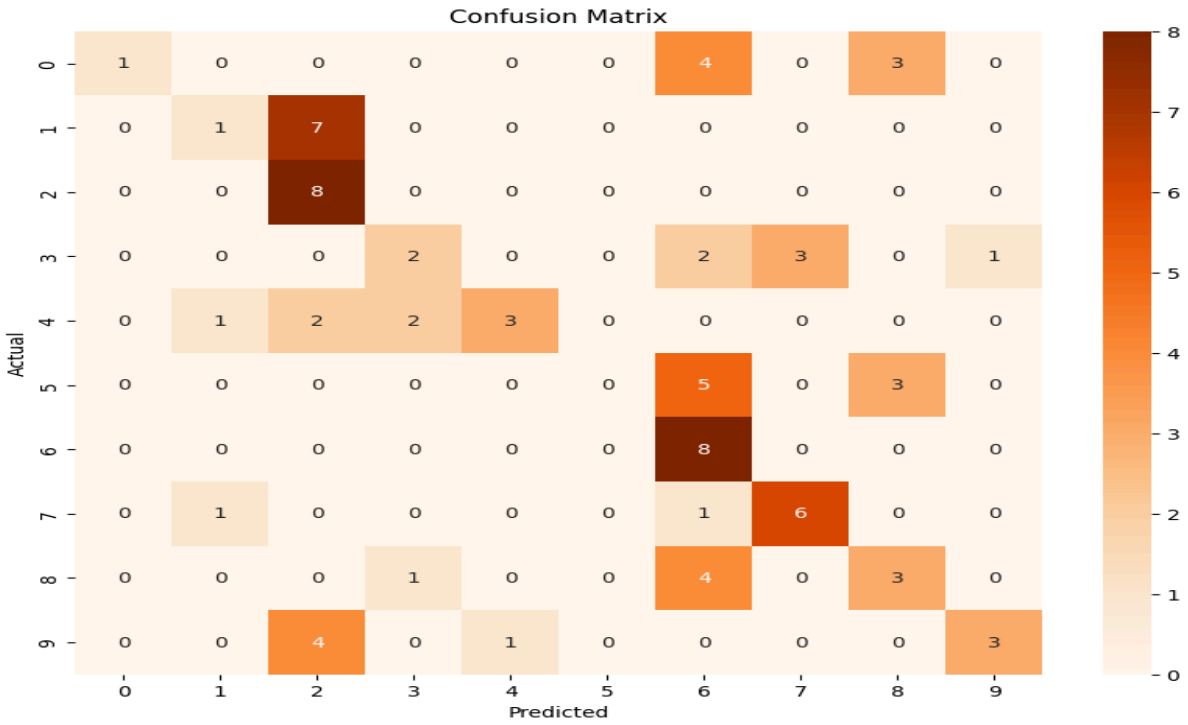
After 100 Epoches results are

Training Loss: 0.4436, Training Accuracy: 98.12%

Validation Loss: 0.2652, Validation Accuracy: 90.31%



Test Loss: 2.2424, Test Accuracy: 43.75%  
F1 SCORE: 0.3822



Total parameters: 484316  
Trainable parameters: 484316  
Non Trainable parameters: 0

## Architecture 2

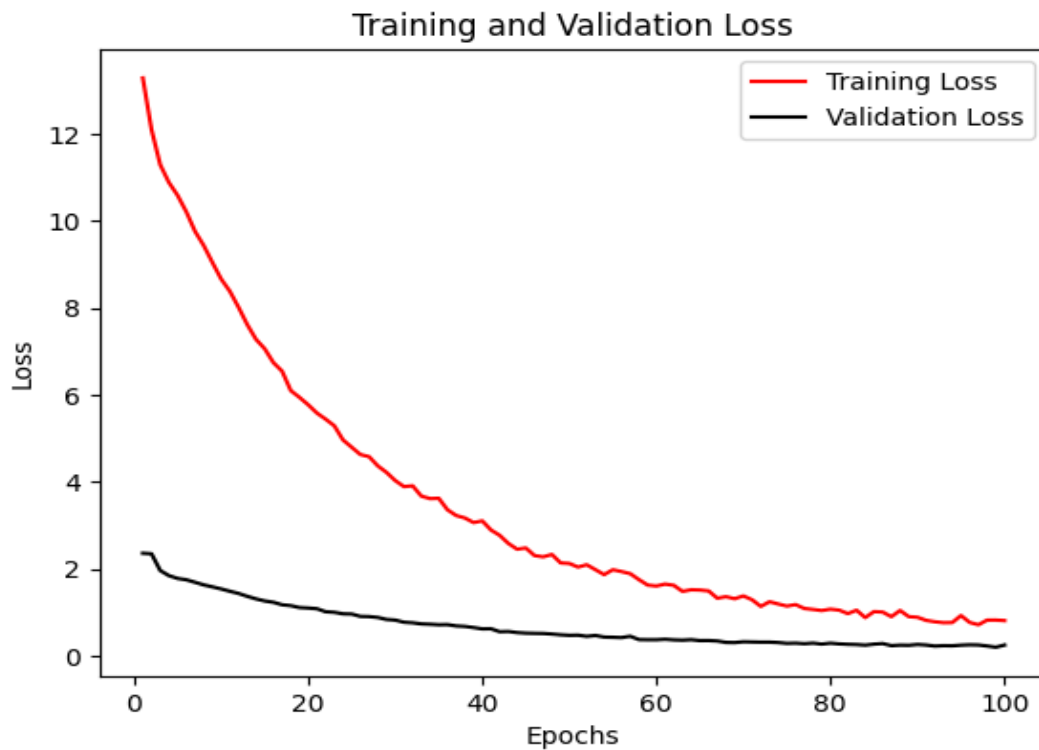
- Transformer with 1 head and 2 encoder blocks.

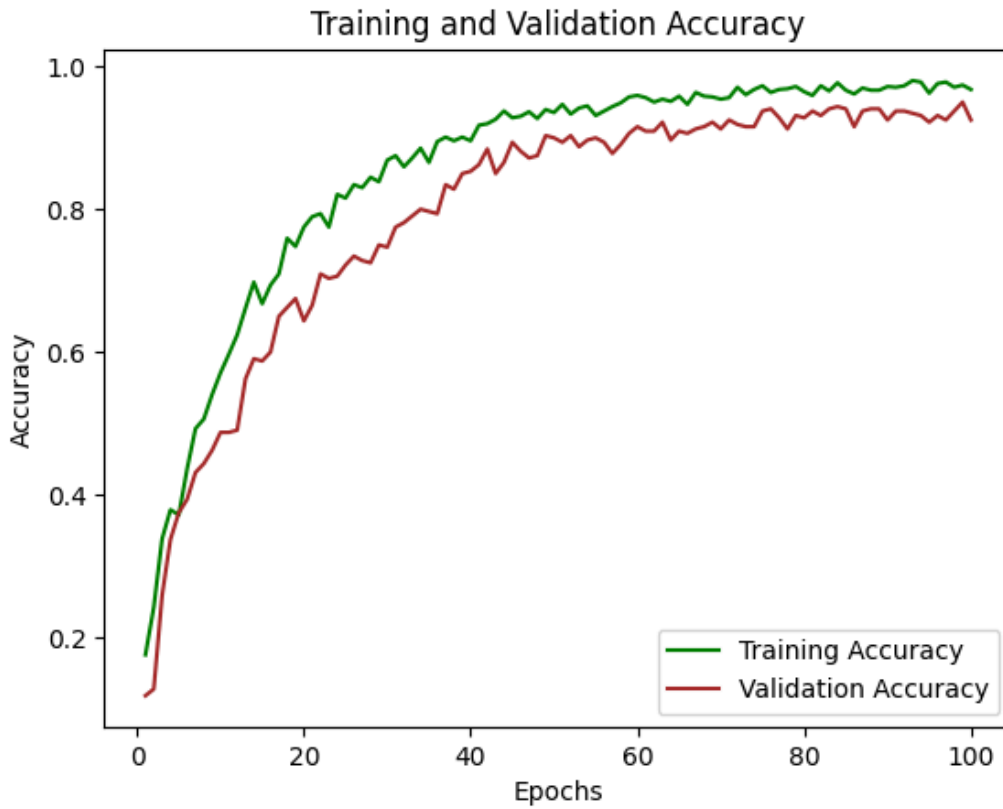
Learning rate = 0.0001, batch size = 40

After 100 epochs

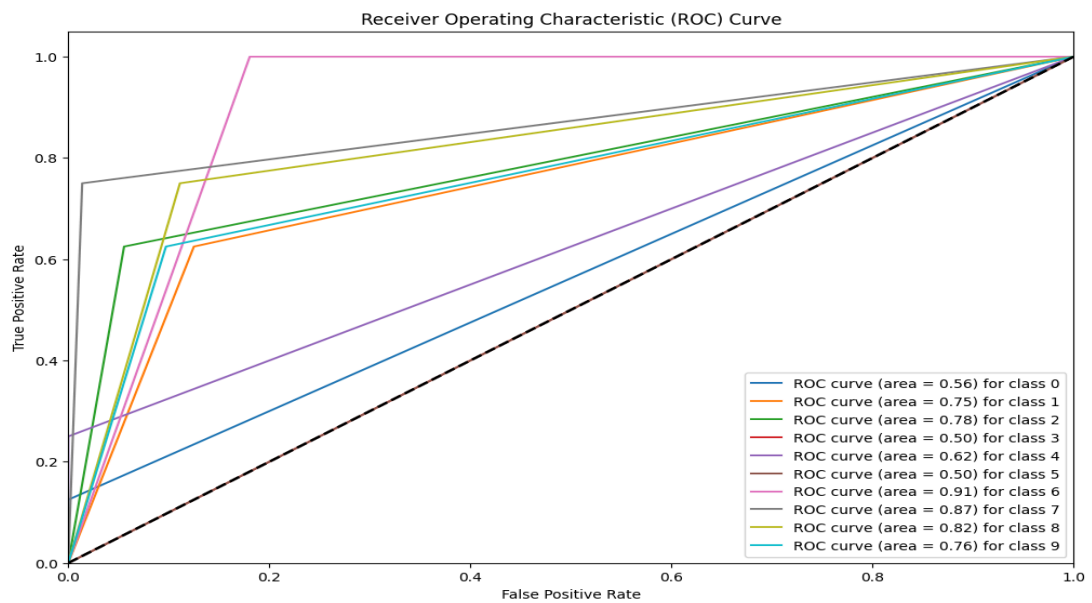
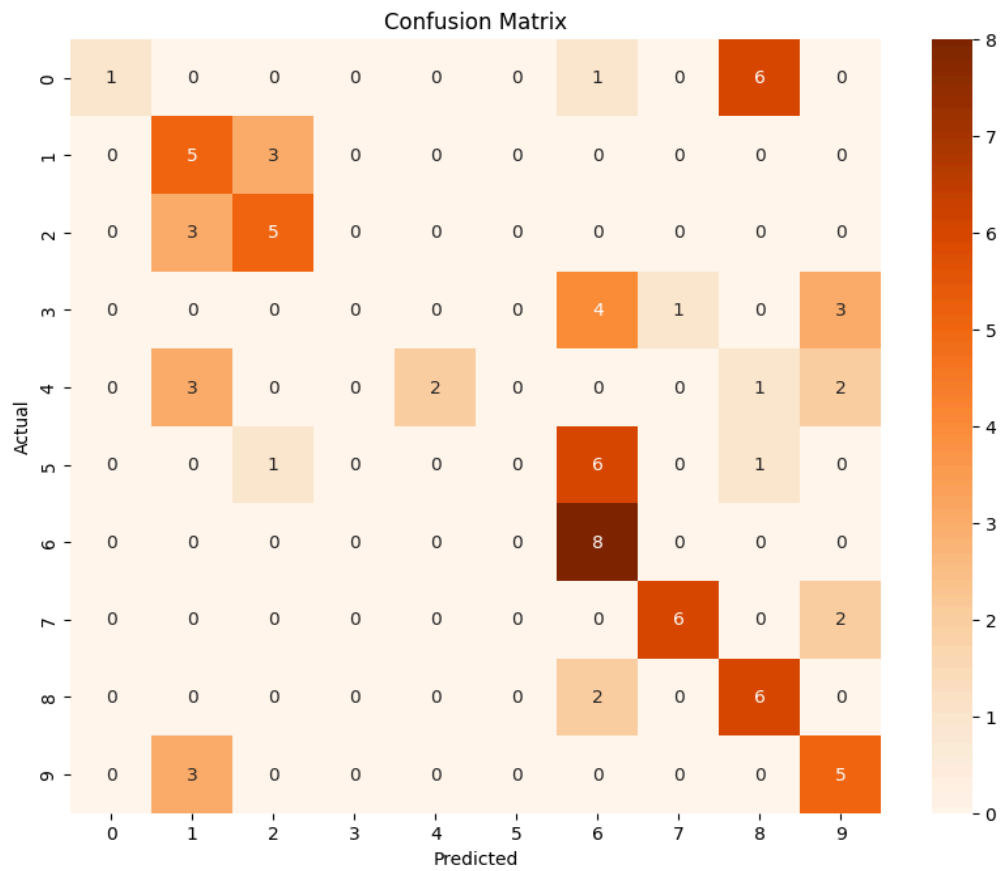
Training Loss: 0.8099, Training Accuracy: 96.77%

Validation Loss: 0.2482, Validation Accuracy: 92.50%





Test Loss: 1.5614, Test Accuracy: 47.50%  
F1 SCORE: 0.4062



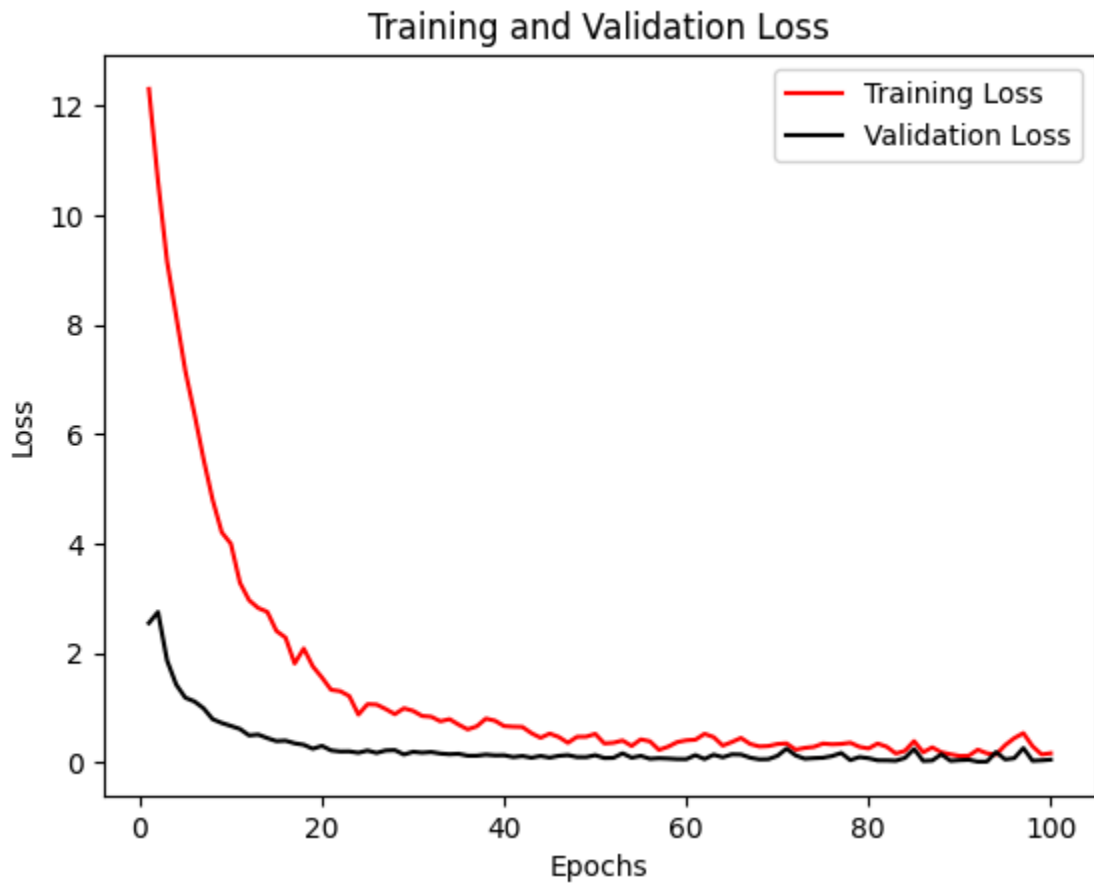
- Transformer with 2 heads and 2 encoder blocks.

Learning rate = 0.0001, batch size = 40

After 100 epochs

Training Loss: 0.1681, Training Accuracy: 99.48%

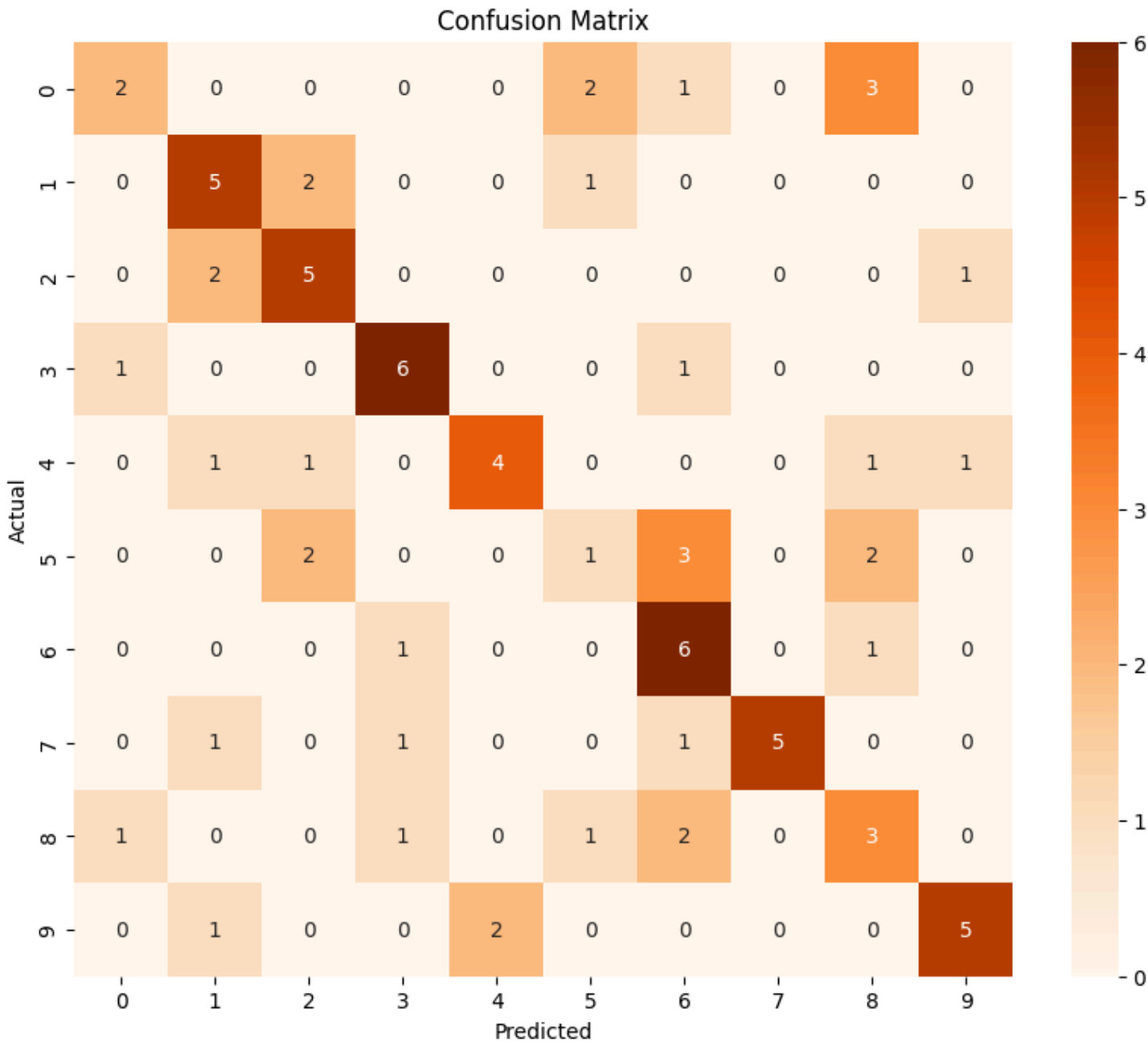
Validation Loss: 0.0513, Validation Accuracy: 98.12%



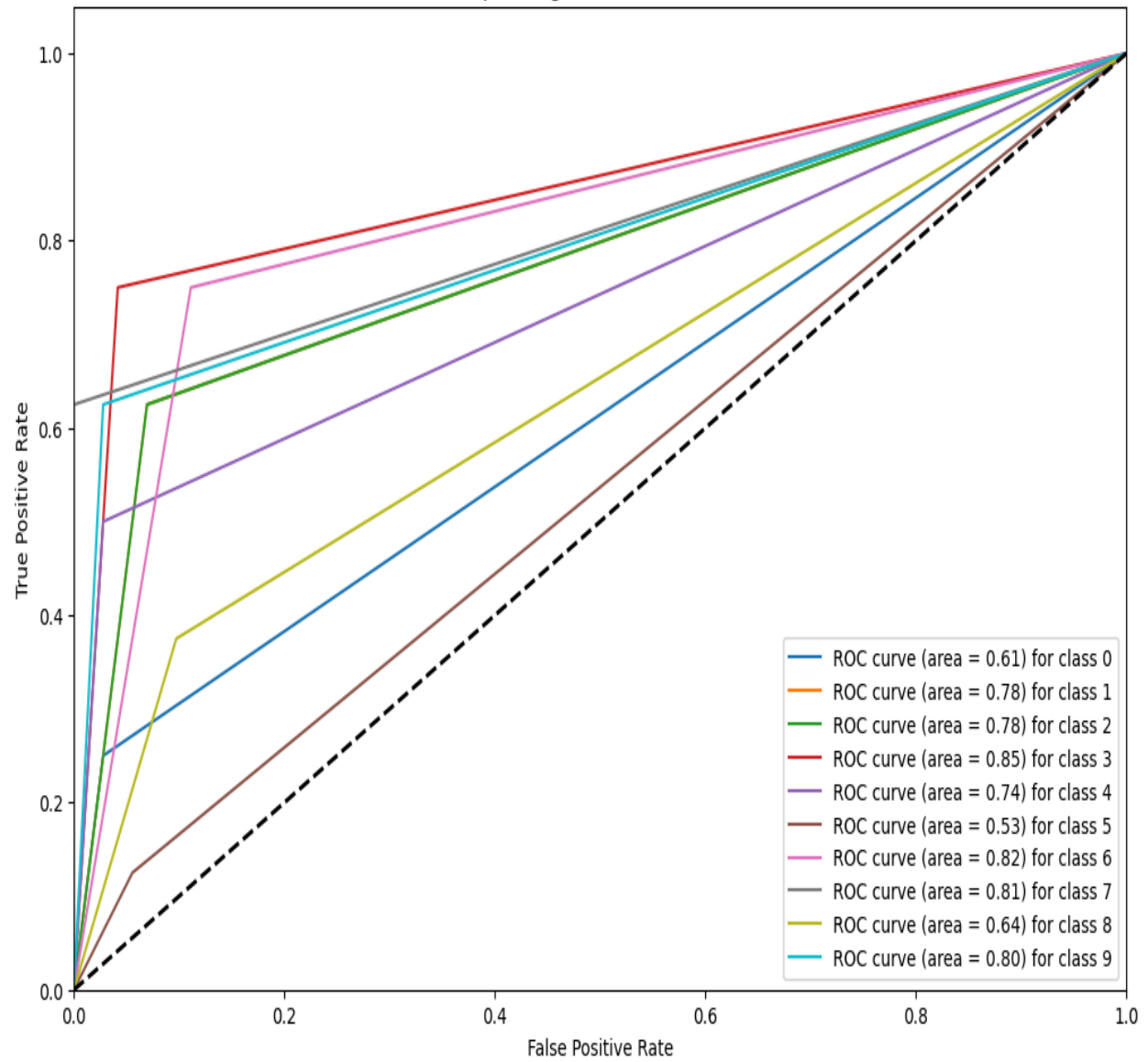




Test Loss: 1.5588, Test Accuracy: 52.50%  
F1 SCORE: 0.5190



Receiver Operating Characteristic (ROC) Curve



- Transformer with 4 heads and 2 encoder blocks.

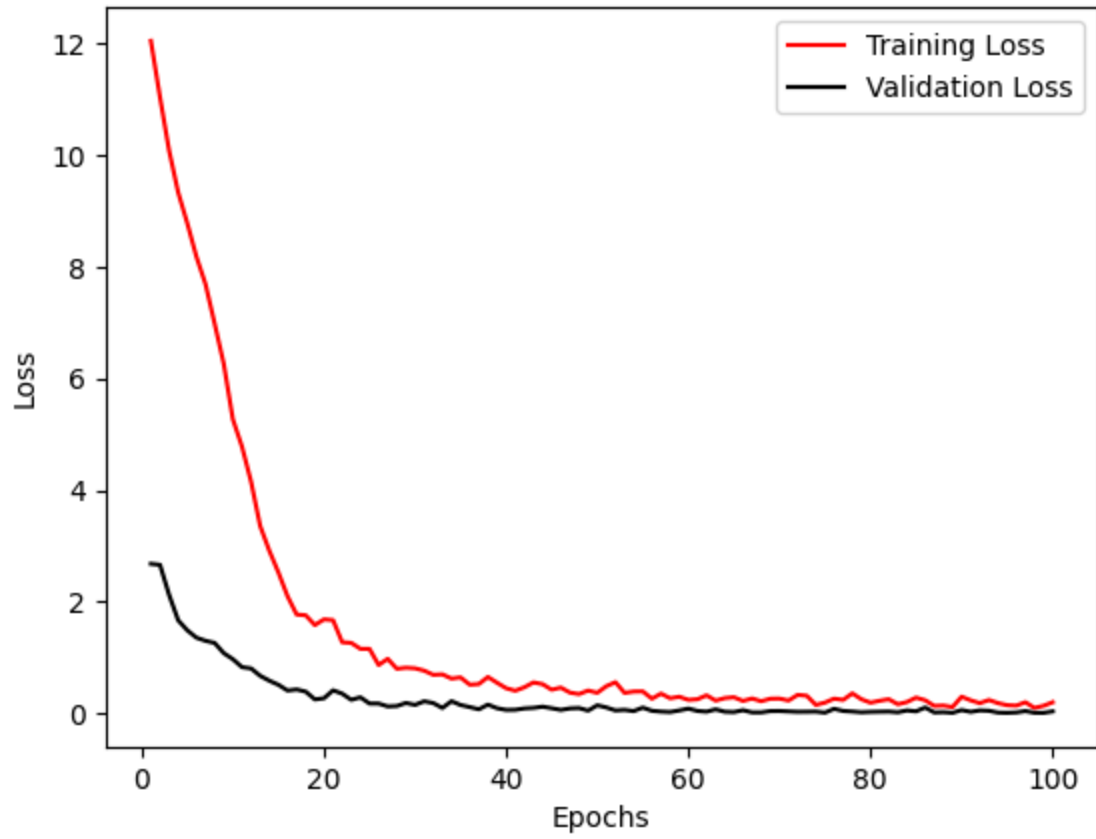
Learning rate = 0.0001, batch size = 40

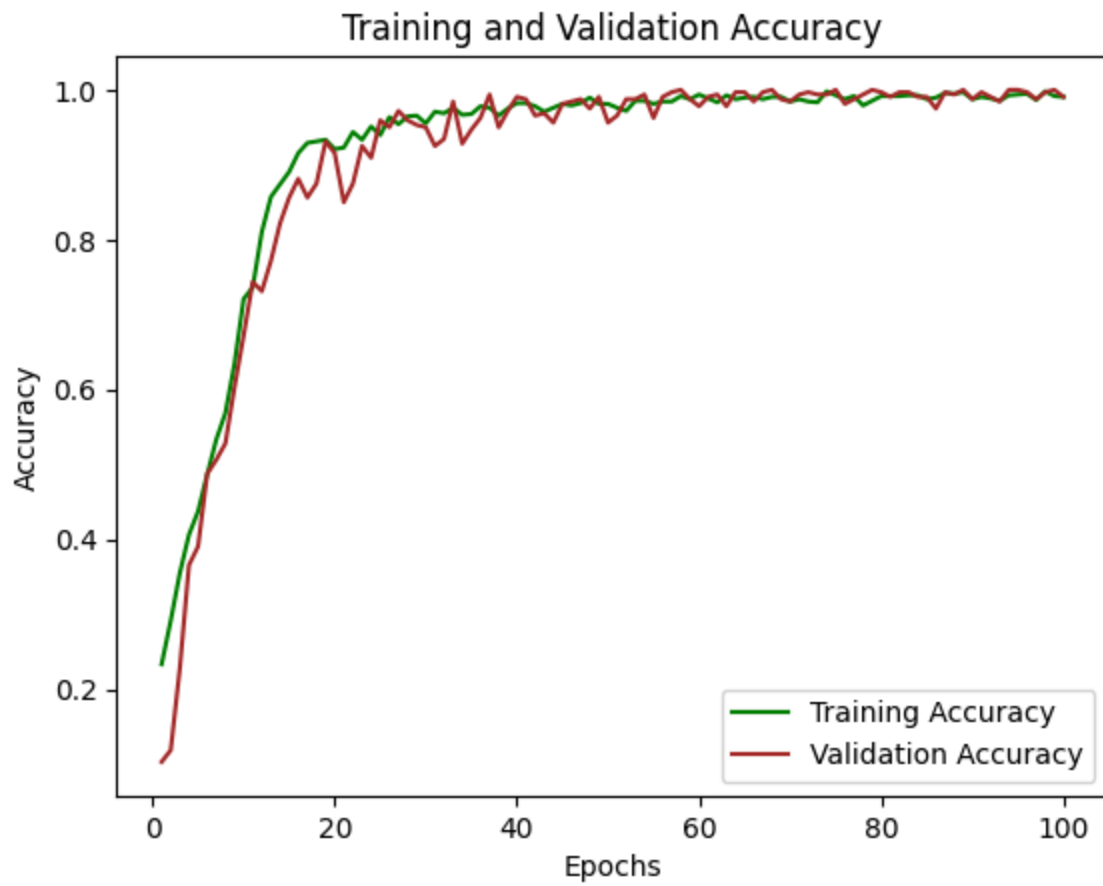
After 100 epochs

Training Loss: 0.1979, Training Accuracy: 99.06%

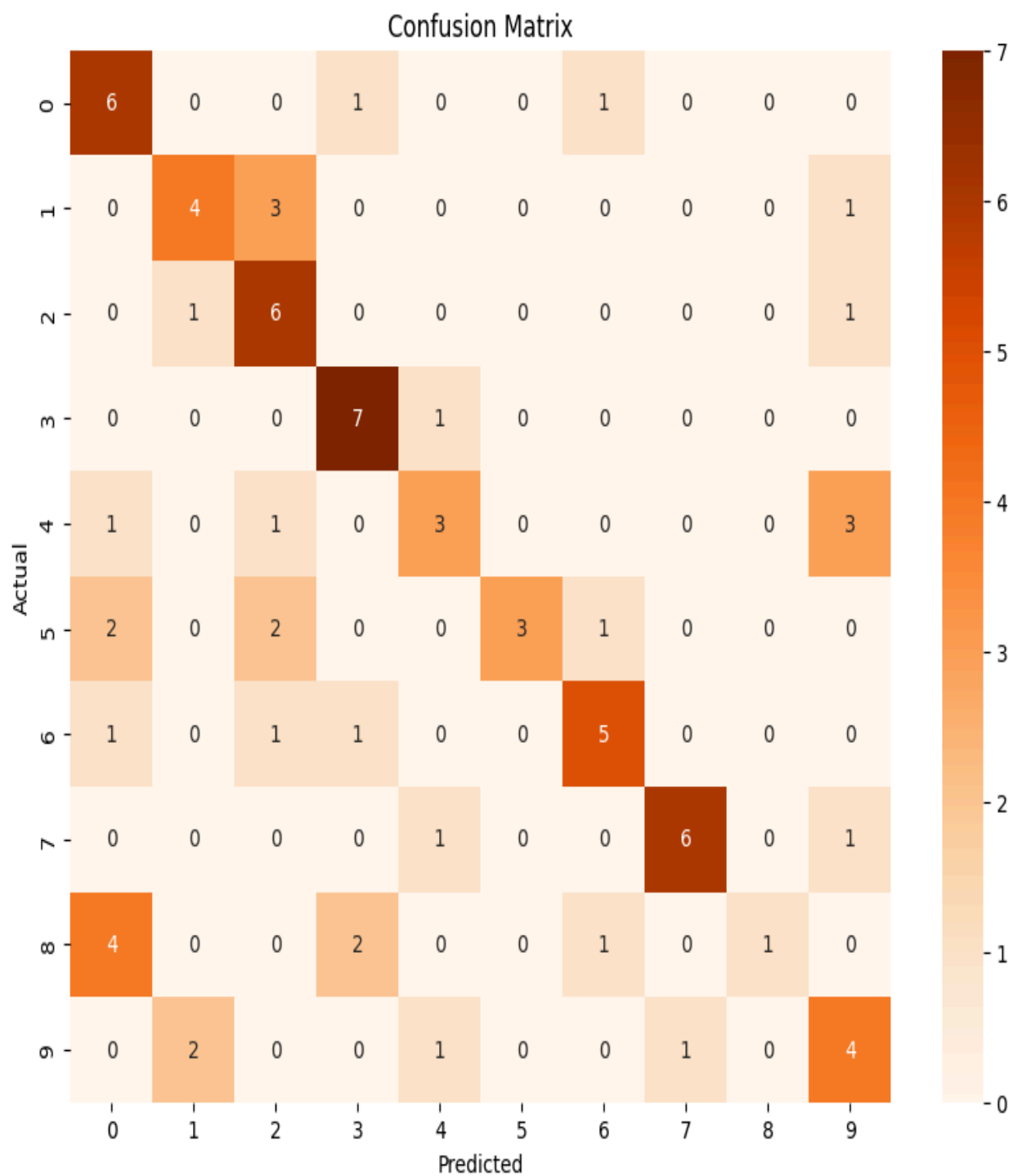
Validation Loss: 0.0341, Validation Accuracy: 99.06%

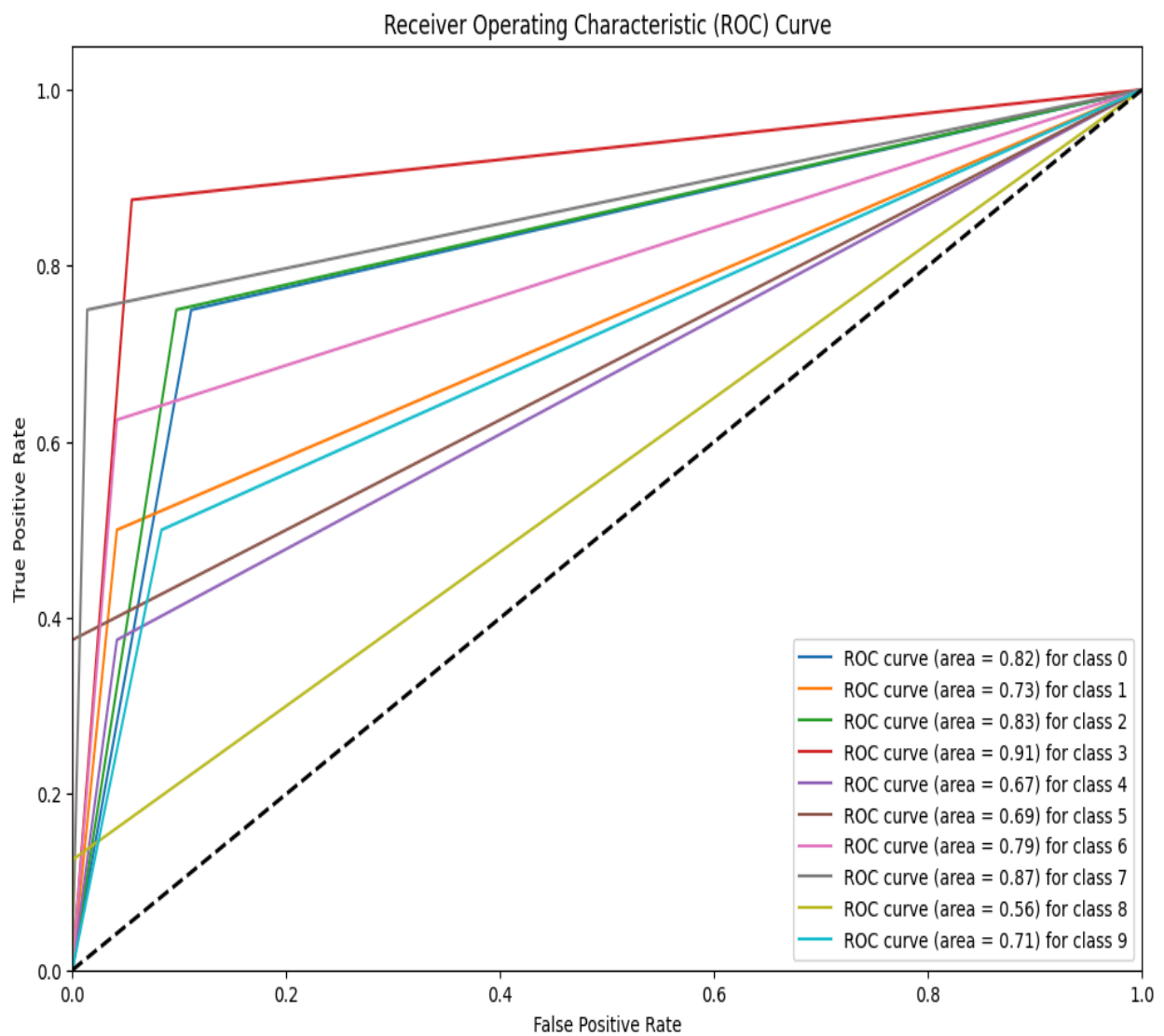
Training and Validation Loss





Test Loss: 1.8032, Test Accuracy: 56.25%  
F1 SCORE: 0.5453





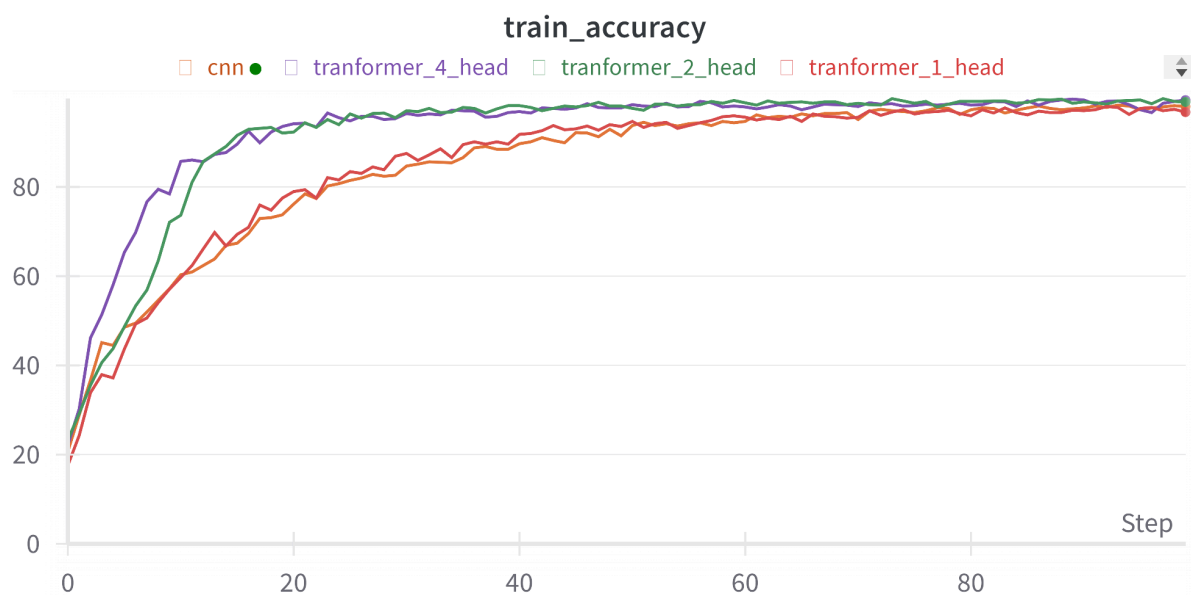


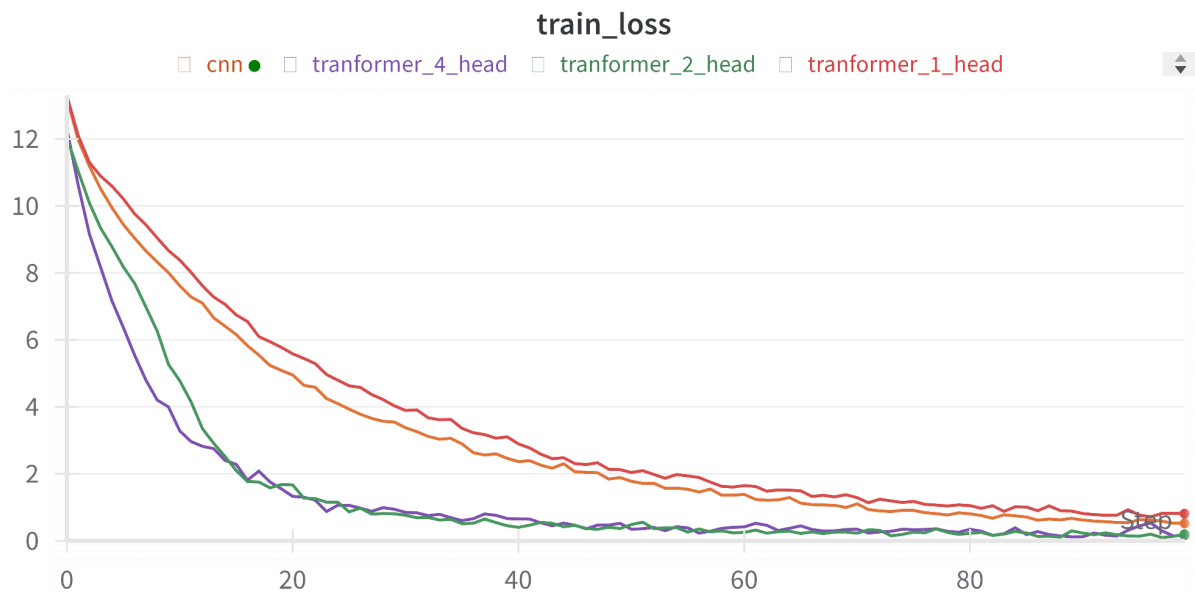
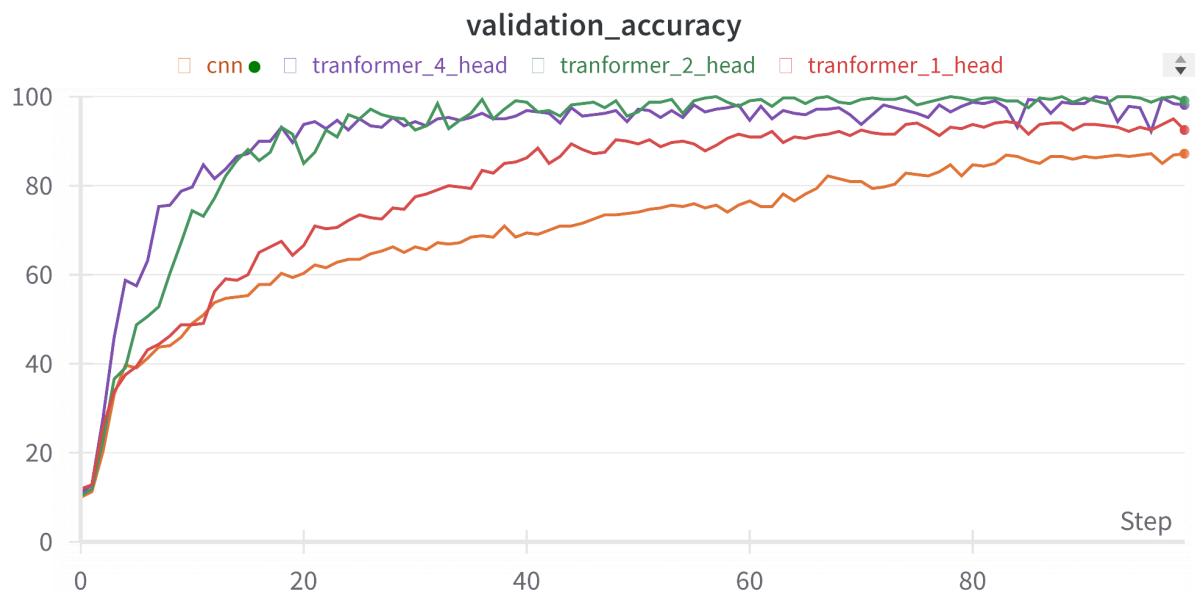
Total\_parameters: 6322984

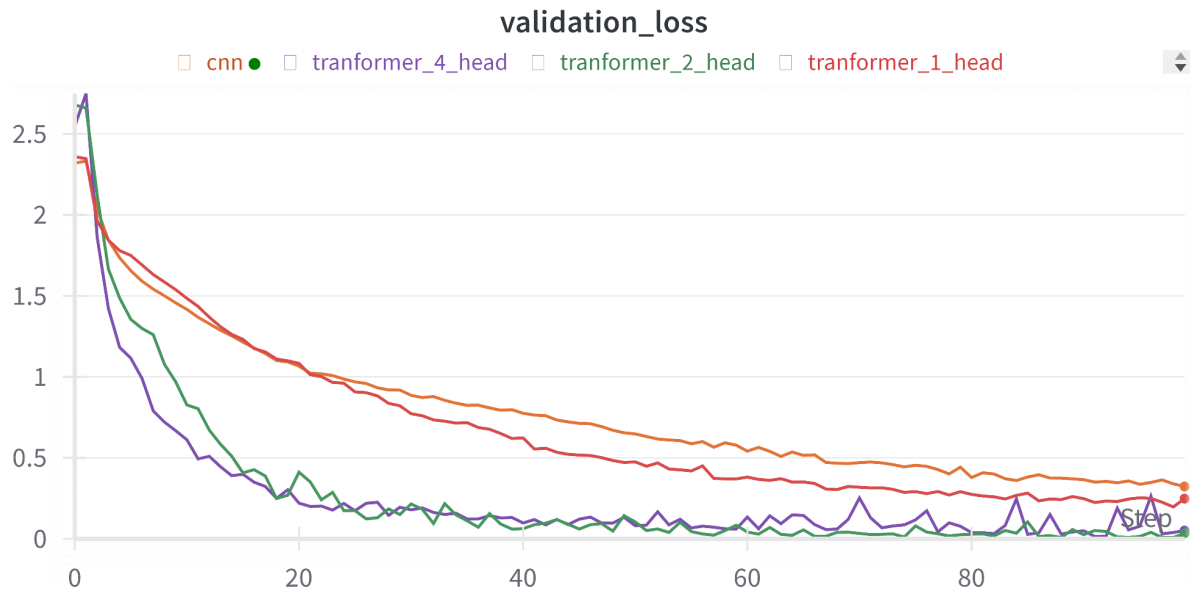
Trainable\_parameters: 6322984

Non\_Trainable\_parameters: 0

Combined loss and accuracy graph







Name	epoch	train_accuracy	validation_accuracy	Test Accuracy	train_loss	validation_loss	Test Loss	Overall F1 Score
cnn	100	98.12%	90.31%	43.75%	0.4436	0.2652	2.2424	0.3822
tranformer_1_head	100	96.77%	92.50%	47.50%	0.8098	0.2482	1.5614	0.4062
tranformer_2_head	100	99.48%	98.12%	52.50%	0.1681	0.0513	1.5588	0.5190
tranformer_4_head	100	98.06%	99.06%	56.25%	0.1979	0.0341	1.8032	0.5453

Train accuracy is almost equal in all the models. There is a overfitting in models and dropout is provided to avoid.

Finally maximum Test accuracy of 56.25 %achieved in transformer with 4 attention head and 2 encoder block.

It is because increasing the number of attention heads, model's capacity of capture complex patterns and relationships within the data increases, which gives better result.