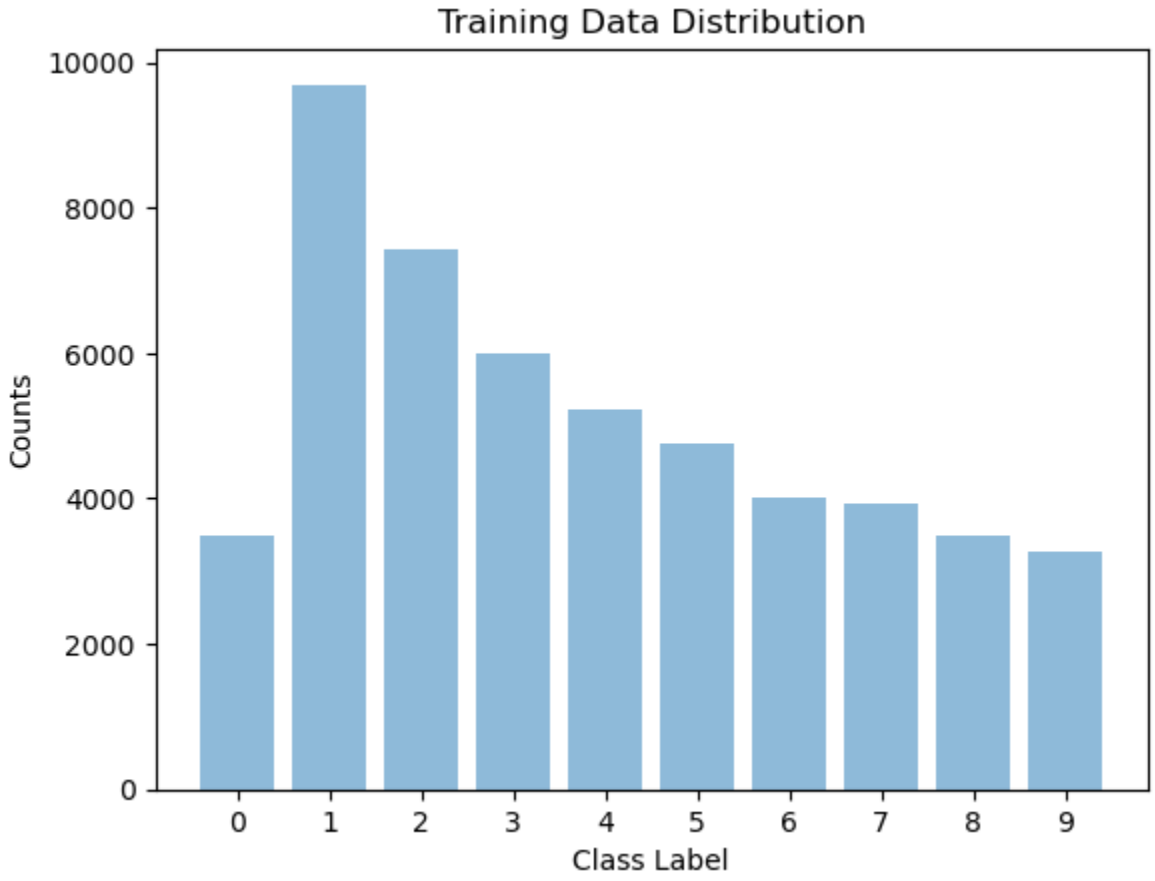
CV Assignment-1

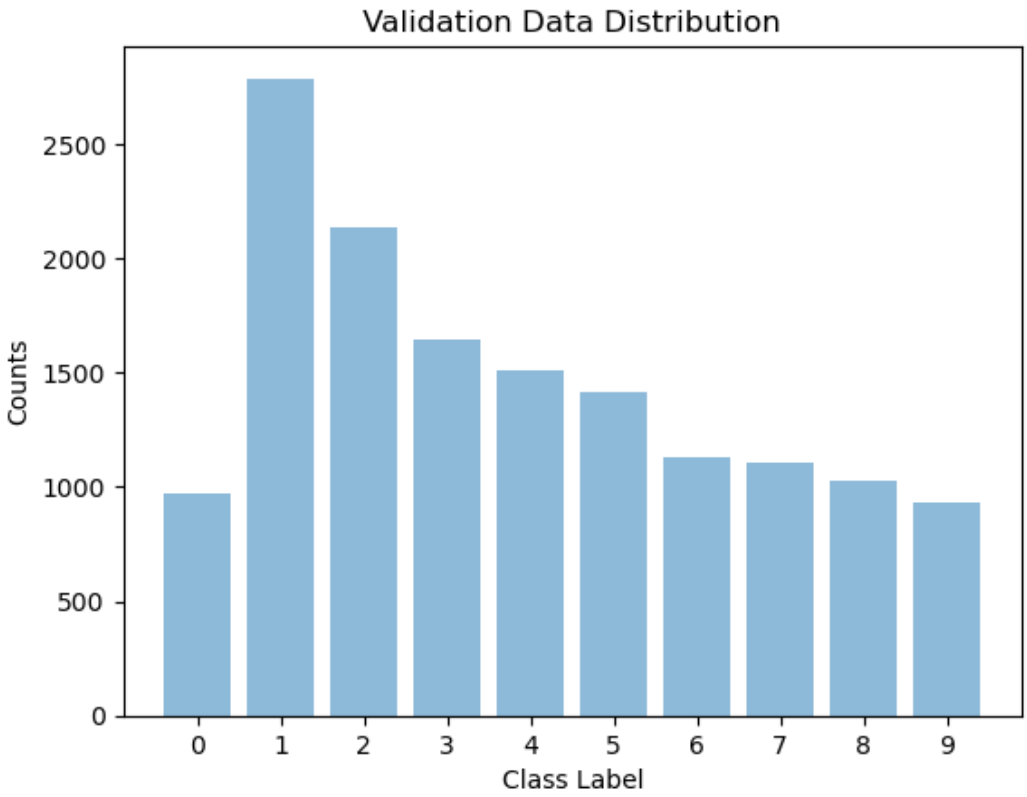
Jaskaran Singh
2020306

Ans 1

I have set class 10 as class 0.

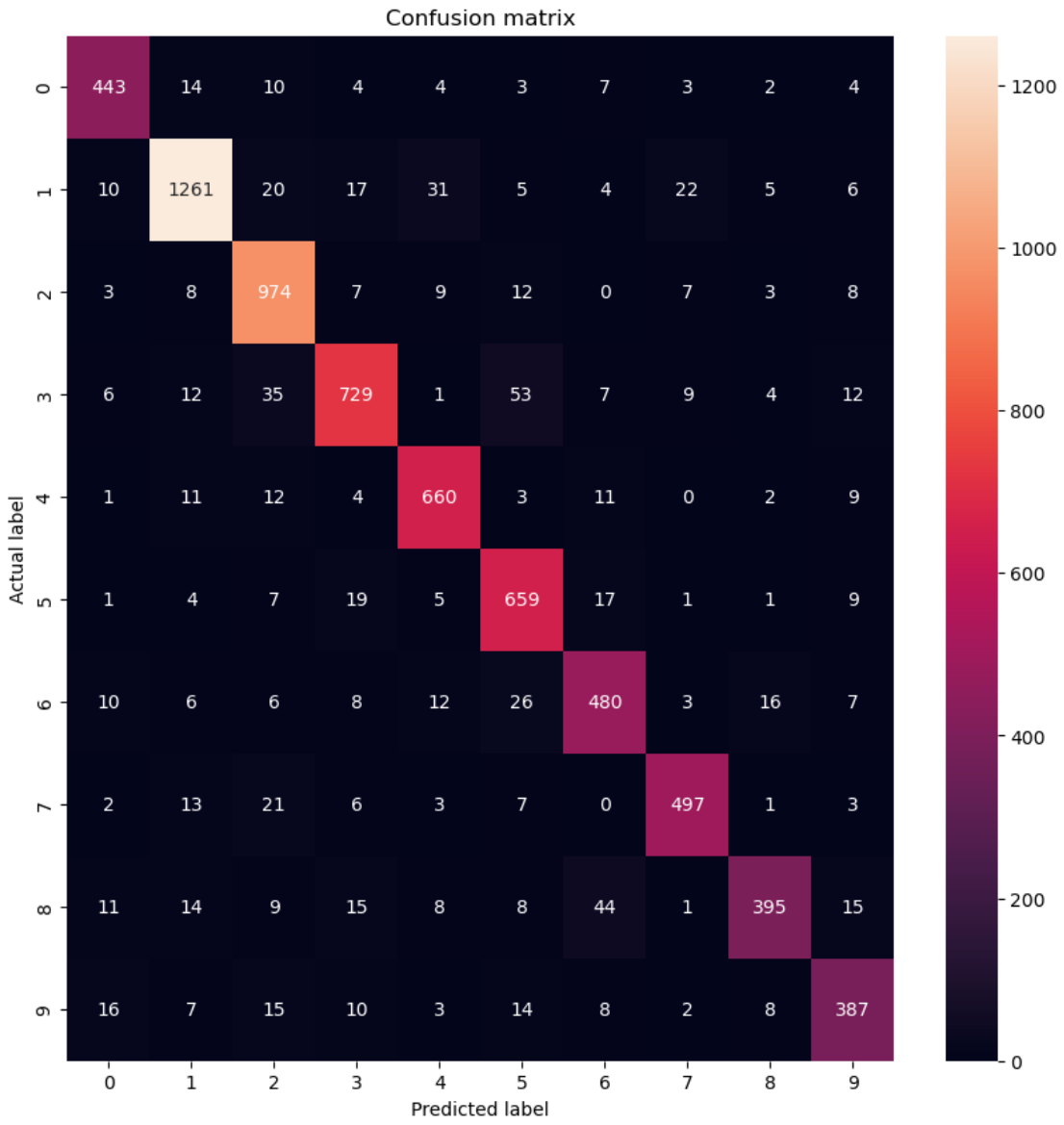
Data distribution visualization graphs:



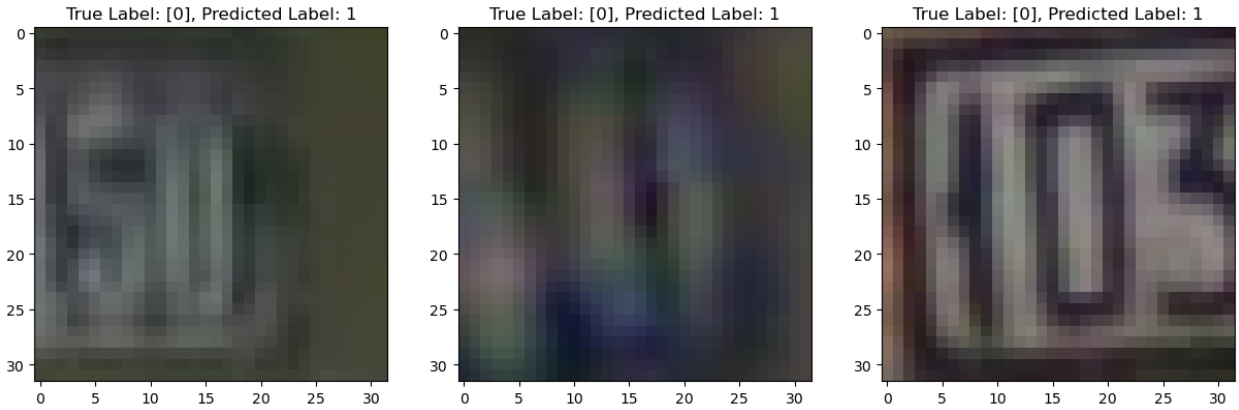


I have created the CNN class with two convolution layers with ReLu and Maxpool and two fully connected layers. The neural network output has ten neurons, each representing a class label.

Confusion matrix:

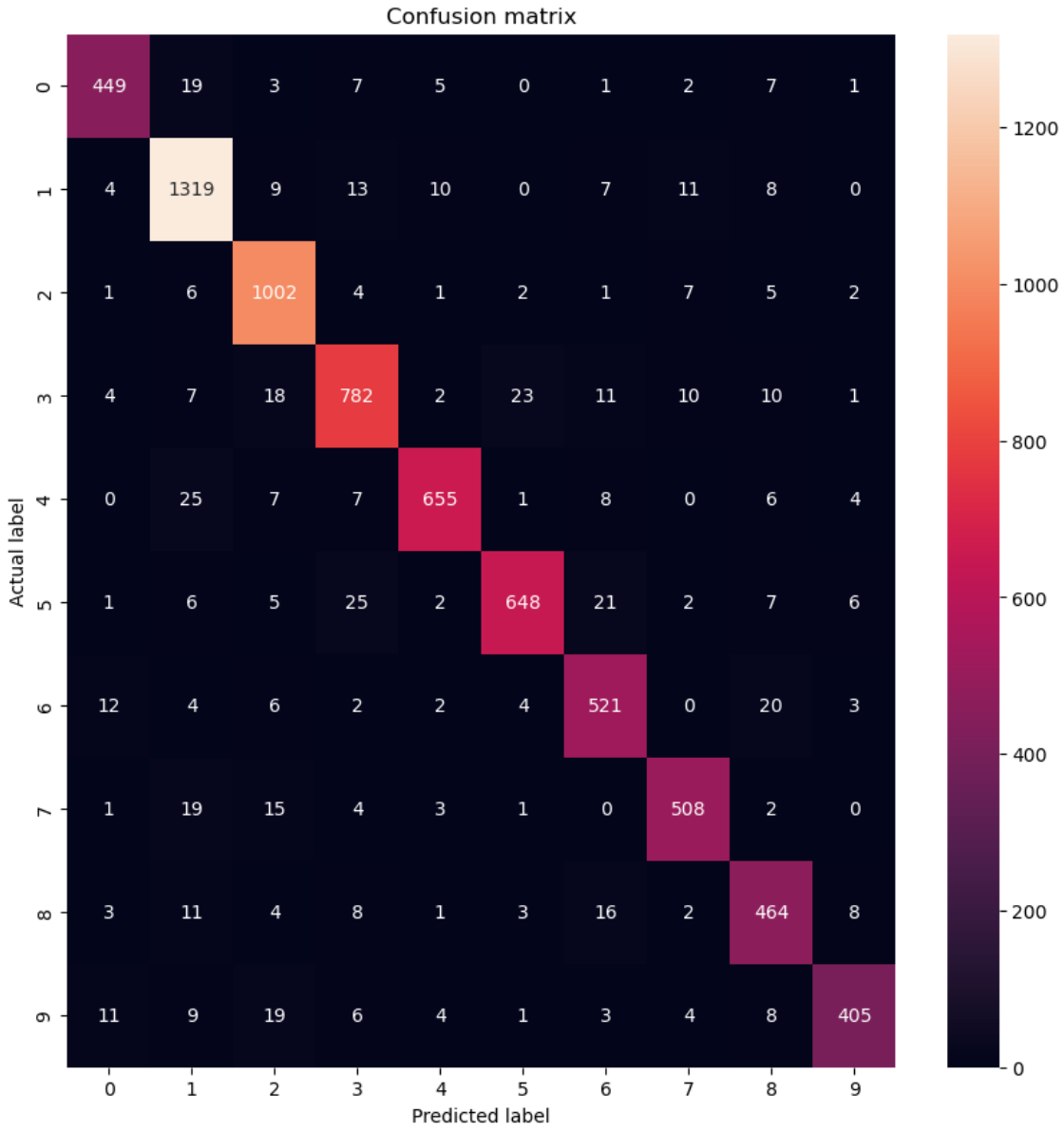


Visualizing three misclassified images for sample class 0:

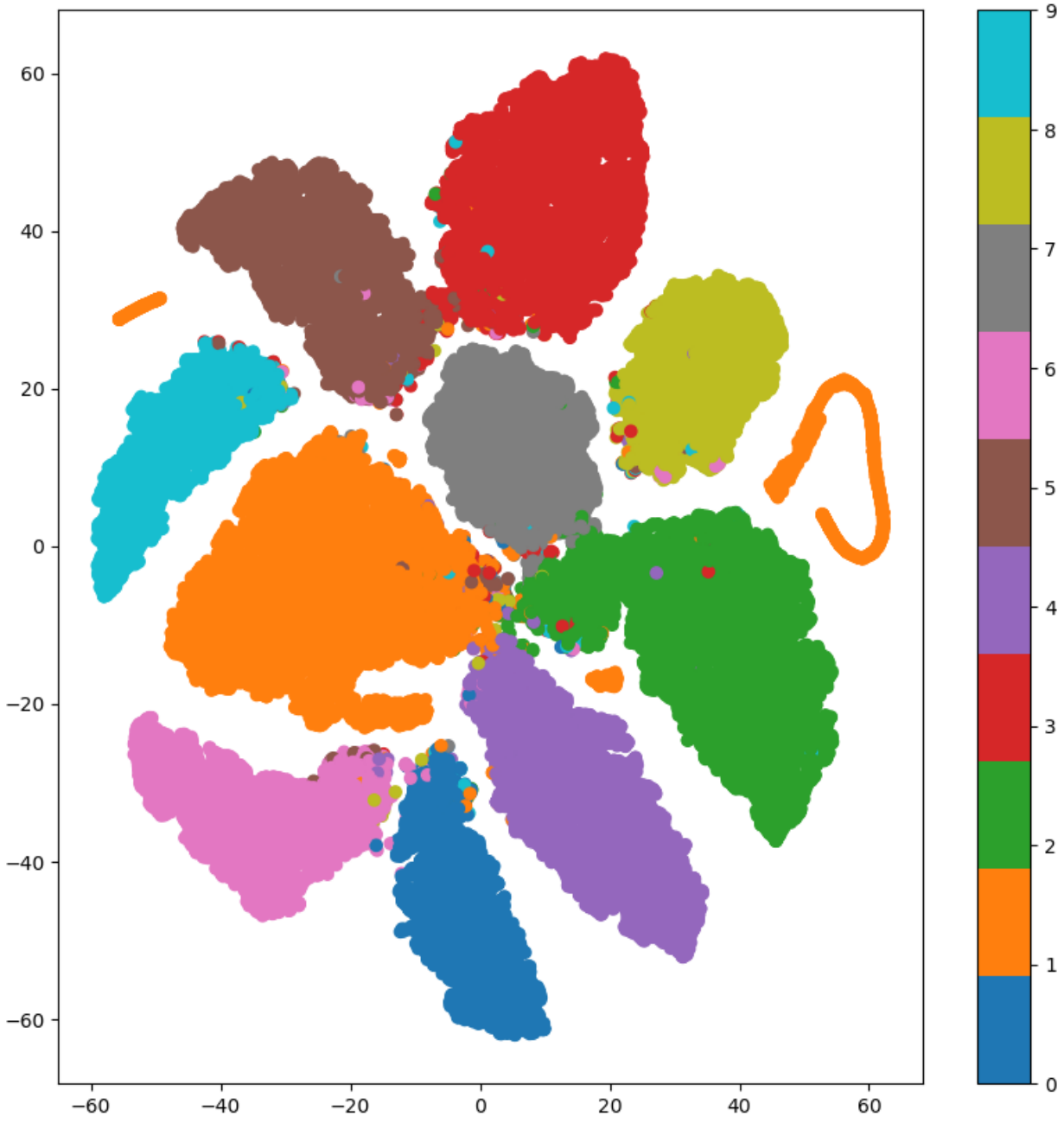


Taking the 3rd image as an example, the image contains 3 numbers, one of which is the number “1” predicted by the model. This is because there are 3 numbers in this testing sample, while the model was only trained on one. Also, the predicted label “1” is the first number in this sample, which might be why this was predicted by the model instead of the true label “0”.
Also, when visualizing misclassified examples from other classes, there is a mismatch because the predicted sample looks closely like the image instead of the ground truth. For example, “3” and “8” are similar, confusing the model. Also, the image size is very small (32x32), and these low-resolution images can easily confuse the model. This can be further seen through the tSNE plots shown below.

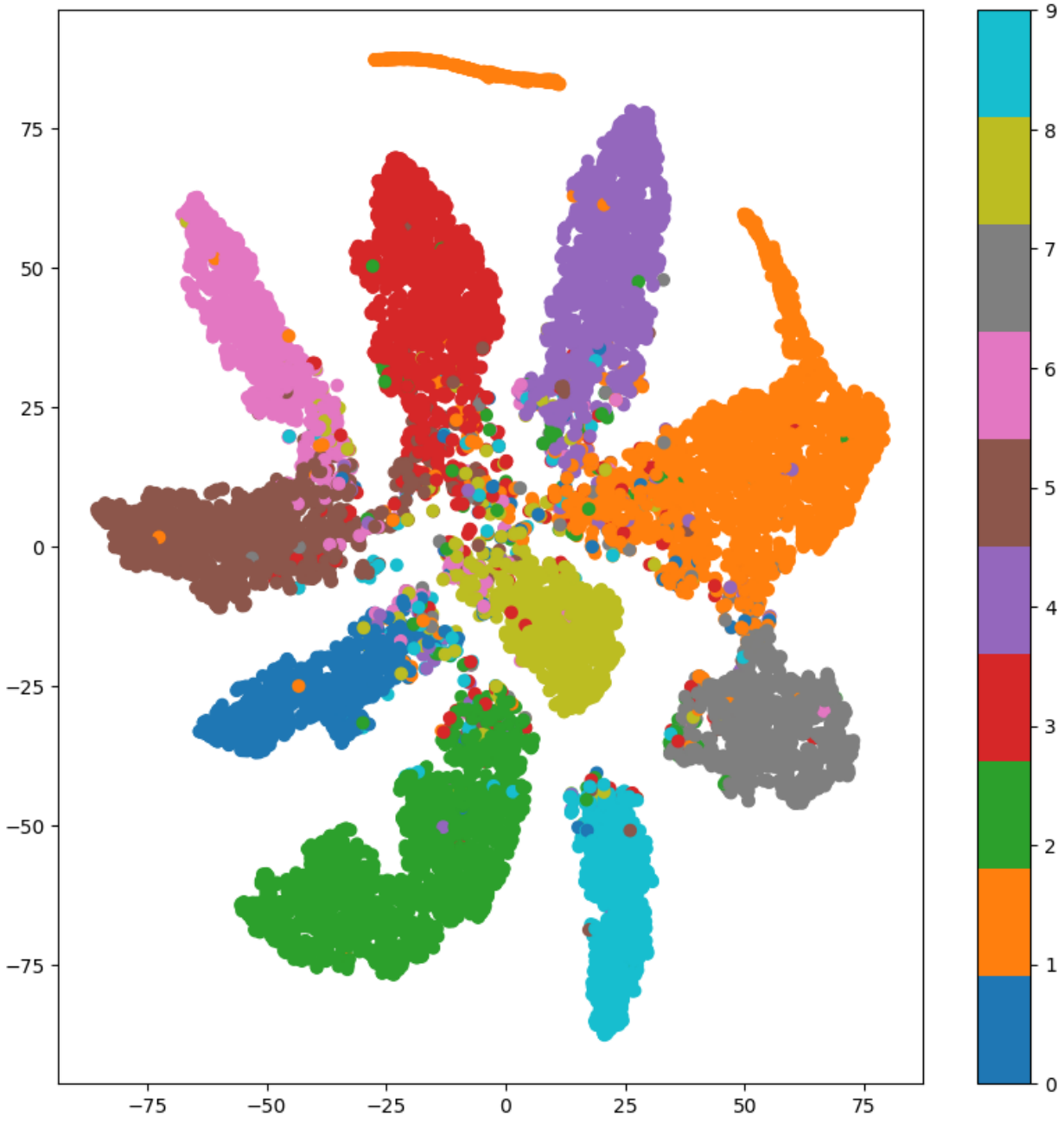
Confusion matrix for Resnet18 model:



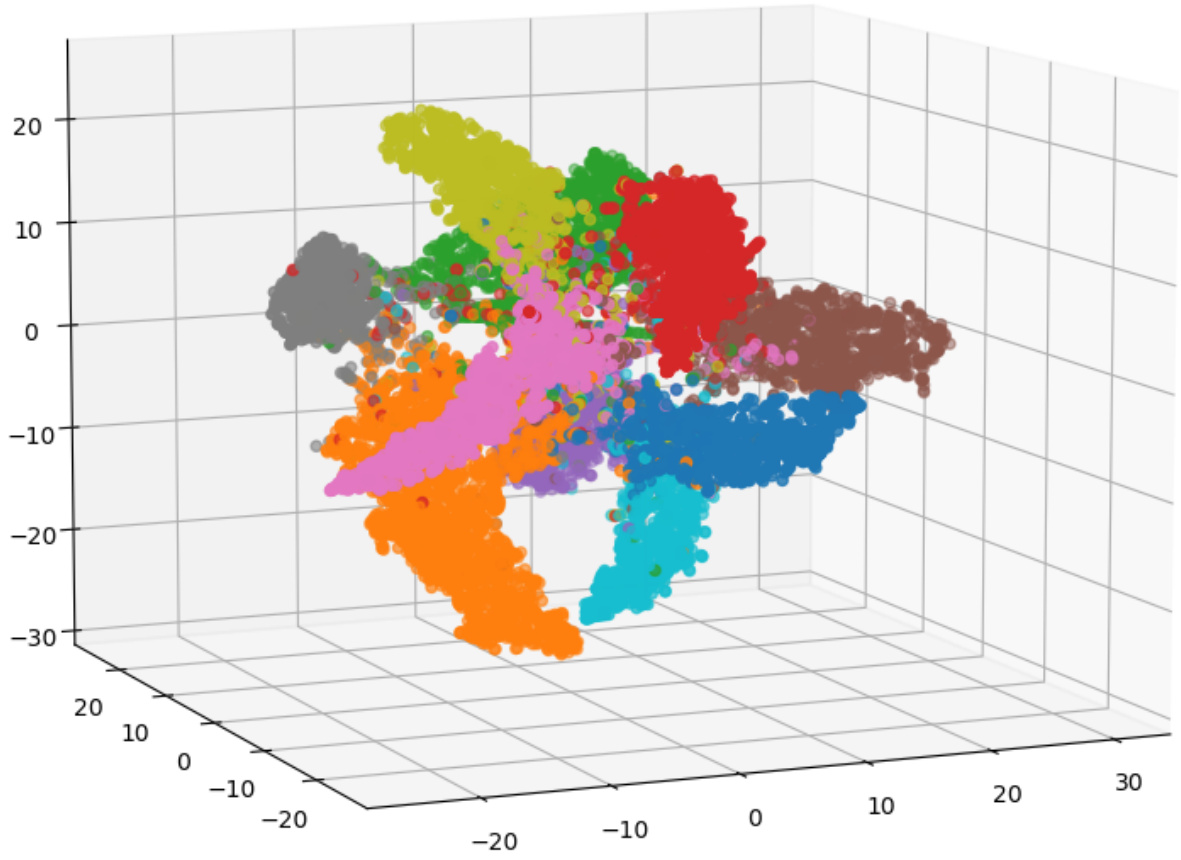
tSNE plot from training data:



tSNE plot from validation data in 2d space:

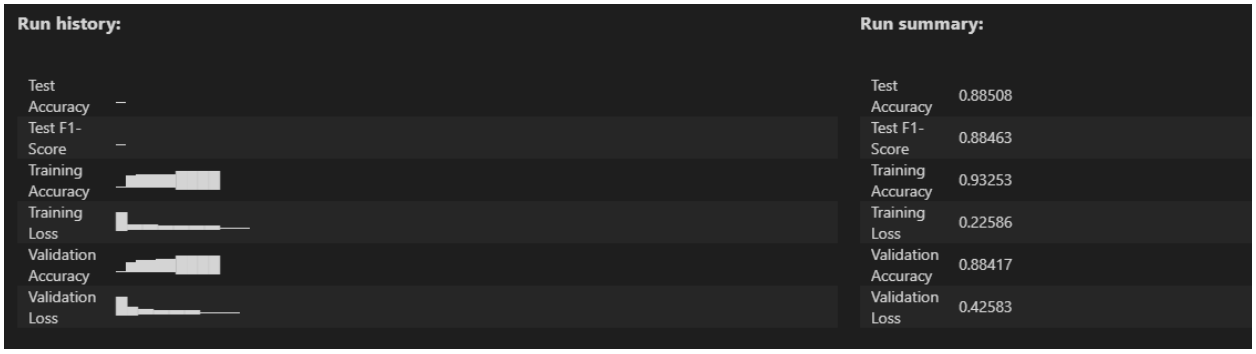


tSNE plot from validation data in 3d space:

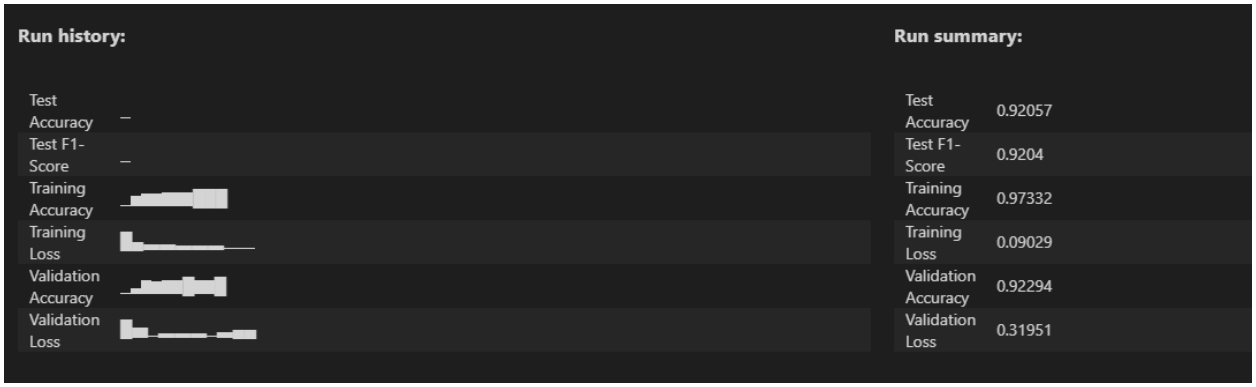


Comparison of the models:

CNN:



ResNet18:

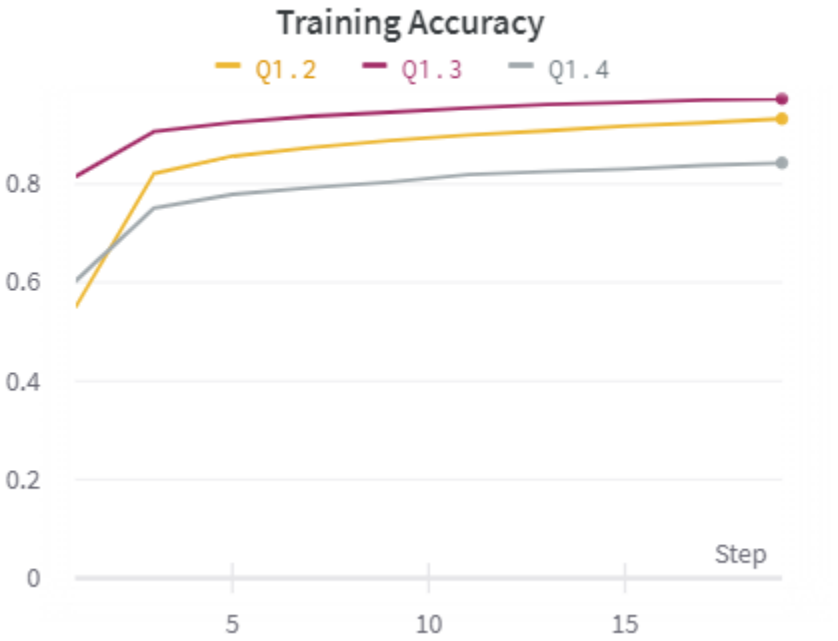


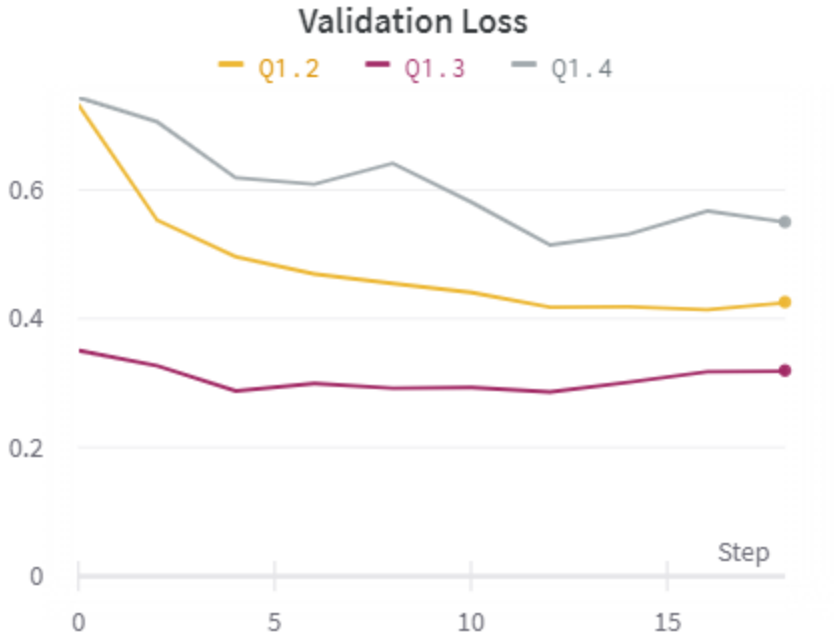
Augmented ResNet18:

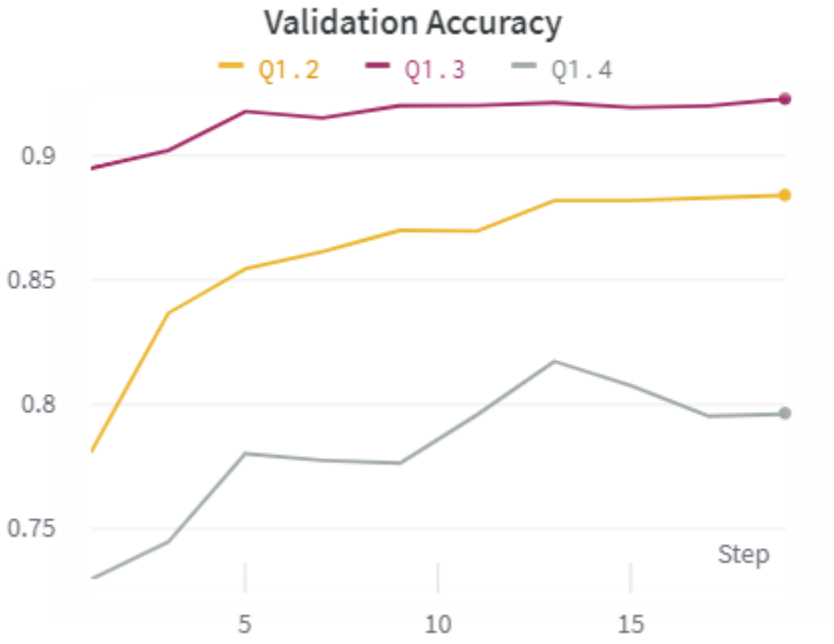
The performance of the models is ReNet18 > Augmented ResNet18 > CNN.
This is because our CNN is a very small model compared to the ResNet18 model, which has many layers. Also, as the Augmented model is trained on a varied dataset, it shows lower scores than the basic ResNet18 model as it overfits the training data. However, if we train the models for a higher number of epochs, we will find that the validation loss of Augmented ResNet is lower than that of base ResNet as it will not overfit. So, in the long run, Augmented ResNet might give better validation and testing results.

WandB plots:


Training Loss
Q1.2
Q1.3
Q1.4
1.2
1
0.8
0.6
0.4
0.2
0
Step
0
5
10
15

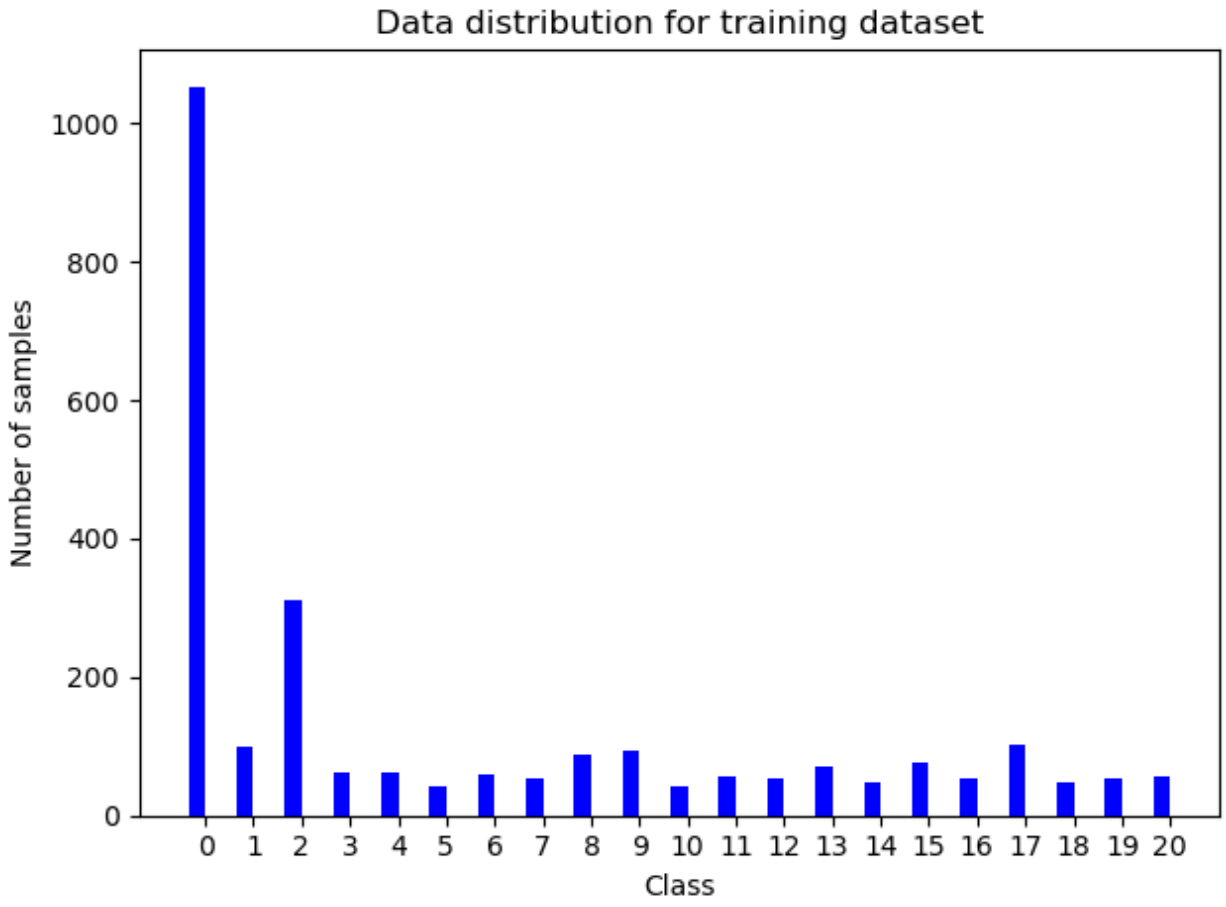

Training Accuracy
Q1.2
Q1.3
Q1.4
0.8
0.6
0.4
0.2
0
Step
5
10
15

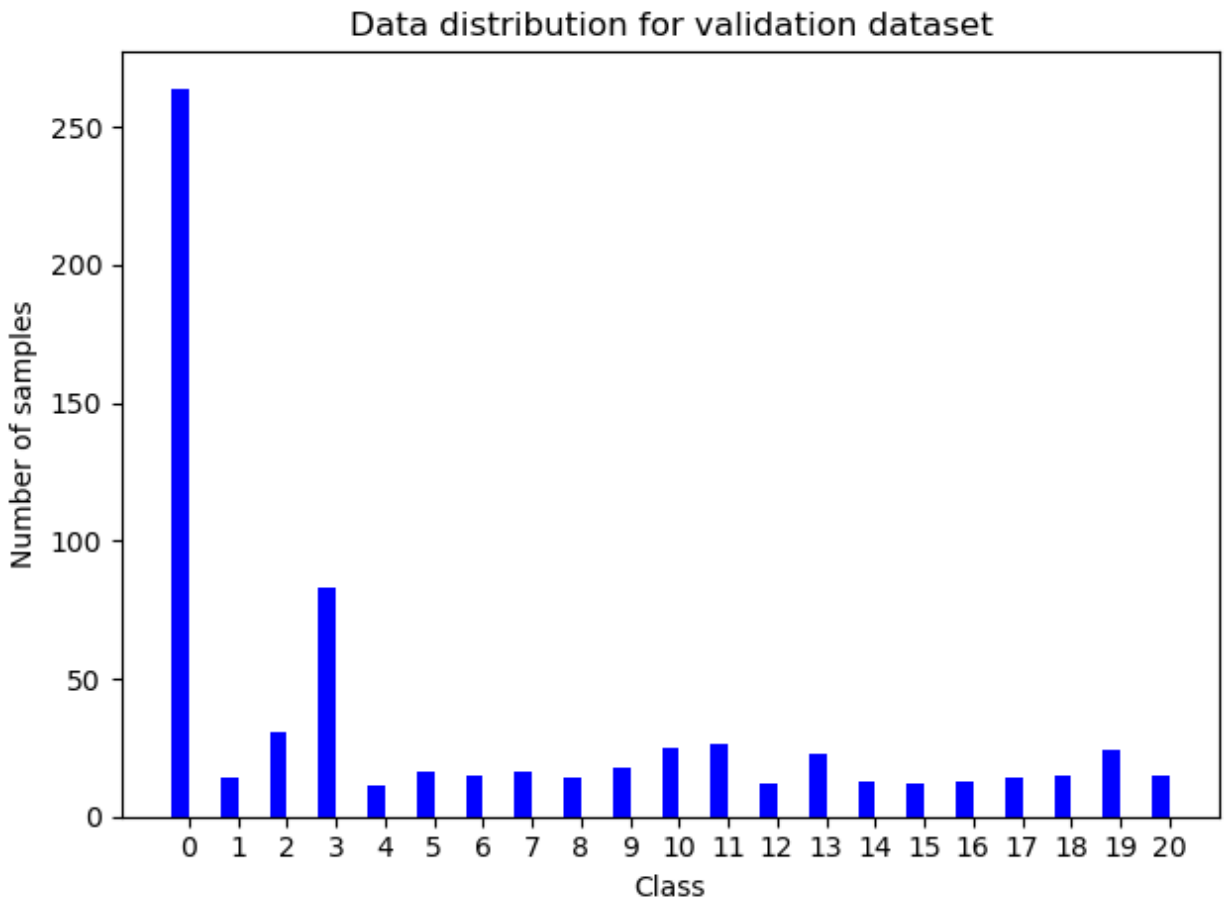

Validation Loss
Q1.2
Q1.3
Q1.4
0.6
0.4
0.2
0
Step
0
5
10
15


Validation Accuracy
Q1.2
Q1.3
Q1.4
0.9
0.85
0.8
0.75
Step
5
10
15

Ans 2

Data distributions:

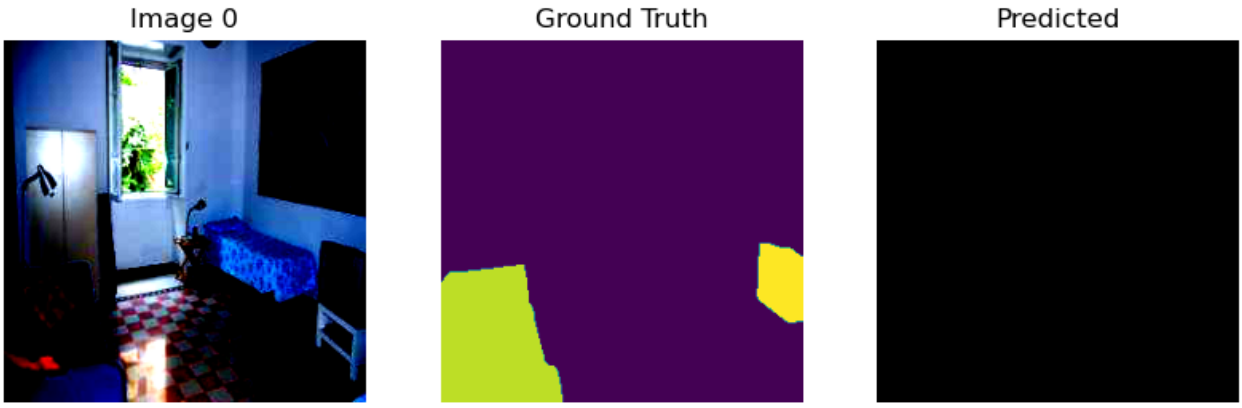




Metrics:

```
Class 0 - Accuracy: 0.9404 - Precision: 0.9481 - Recall: 0.9756 - F1-Score: 0.9616 - IoU: 0.9261
Class 1 - Accuracy: 0.9984 - Precision: 0.8245 - Recall: 0.9431 - F1-Score: 0.8798 - IoU: 0.7855
Class 2 - Accuracy: 0.9987 - Precision: 0.5633 - Recall: 0.5531 - F1-Score: 0.5581 - IoU: 0.3871
Class 3 - Accuracy: 0.9978 - Precision: 0.8796 - Recall: 0.7509 - F1-Score: 0.8102 - IoU: 0.6809
Class 4 - Accuracy: 0.9982 - Precision: 0.7892 - Recall: 0.7489 - F1-Score: 0.7685 - IoU: 0.6241
Class 5 - Accuracy: 0.9968 - Precision: 0.8448 - Recall: 0.6701 - F1-Score: 0.7474 - IoU: 0.5966
Class 6 - Accuracy: 0.9982 - Precision: 0.9635 - Recall: 0.9313 - F1-Score: 0.9471 - IoU: 0.8996
Class 7 - Accuracy: 0.9981 - Precision: 0.9264 - Recall: 0.8906 - F1-Score: 0.9082 - IoU: 0.8318
Class 8 - Accuracy: 0.9960 - Precision: 0.8932 - Recall: 0.9429 - F1-Score: 0.9173 - IoU: 0.8473
Class 9 - Accuracy: 0.9885 - Precision: 0.7965 - Recall: 0.3399 - F1-Score: 0.4765 - IoU: 0.3127
Class 10 - Accuracy: 0.9987 - Precision: 0.8964 - Recall: 0.8089 - F1-Score: 0.8504 - IoU: 0.7398
Class 11 - Accuracy: 0.9902 - Precision: 0.9281 - Recall: 0.6878 - F1-Score: 0.7901 - IoU: 0.6530
Class 12 - Accuracy: 0.9966 - Precision: 0.8096 - Recall: 0.8653 - F1-Score: 0.8366 - IoU: 0.7190
Class 13 - Accuracy: 0.9988 - Precision: 0.8122 - Recall: 0.8221 - F1-Score: 0.8171 - IoU: 0.6908
Class 14 - Accuracy: 0.9976 - Precision: 0.7933 - Recall: 0.8221 - F1-Score: 0.8074 - IoU: 0.6771
Class 15 - Accuracy: 0.9914 - Precision: 0.8987 - Recall: 0.8965 - F1-Score: 0.8976 - IoU: 0.8142
Class 16 - Accuracy: 0.9983 - Precision: 0.7960 - Recall: 0.8703 - F1-Score: 0.8315 - IoU: 0.7116
Class 17 - Accuracy: 0.9984 - Precision: 0.9408 - Recall: 0.9222 - F1-Score: 0.9314 - IoU: 0.8716
Class 18 - Accuracy: 0.9954 - Precision: 0.9402 - Recall: 0.6762 - F1-Score: 0.7867 - IoU: 0.6484
Class 19 - Accuracy: 0.9983 - Precision: 0.9182 - Recall: 0.9788 - F1-Score: 0.9476 - IoU: 0.9003
Class 20 - Accuracy: 0.9967 - Precision: 0.9999 - Recall: 0.3105 - F1-Score: 0.4738 - IoU: 0.3105
Average precision score: 0.2642
```

Sample prediction:

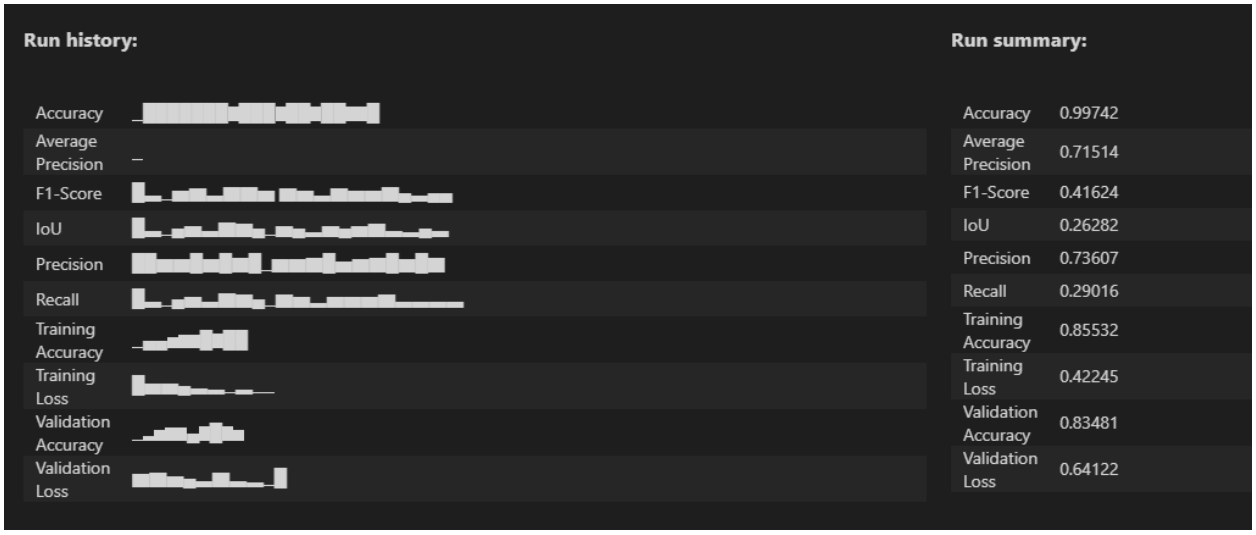


Here, the model gave no prediction. This might be because the model cannot identify the objects in the image. This is due to the surroundings of the object, due to which it is not being classified.

Performance:

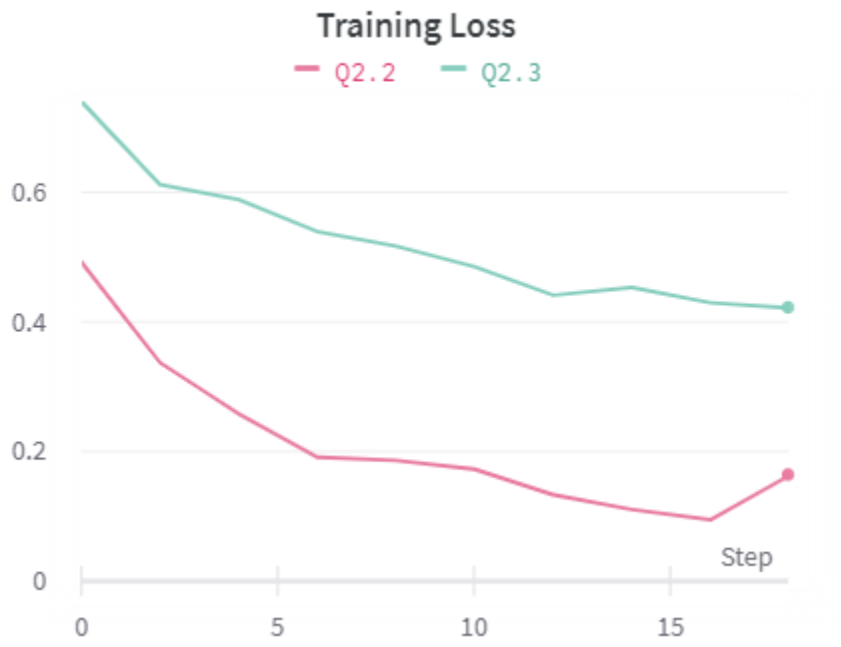
ResNet50:

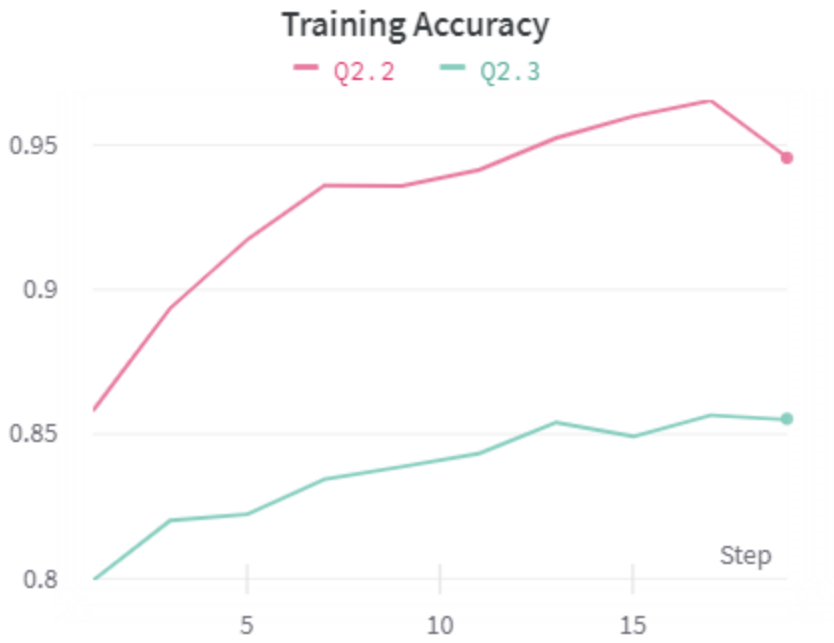
Augmented ResNet50:

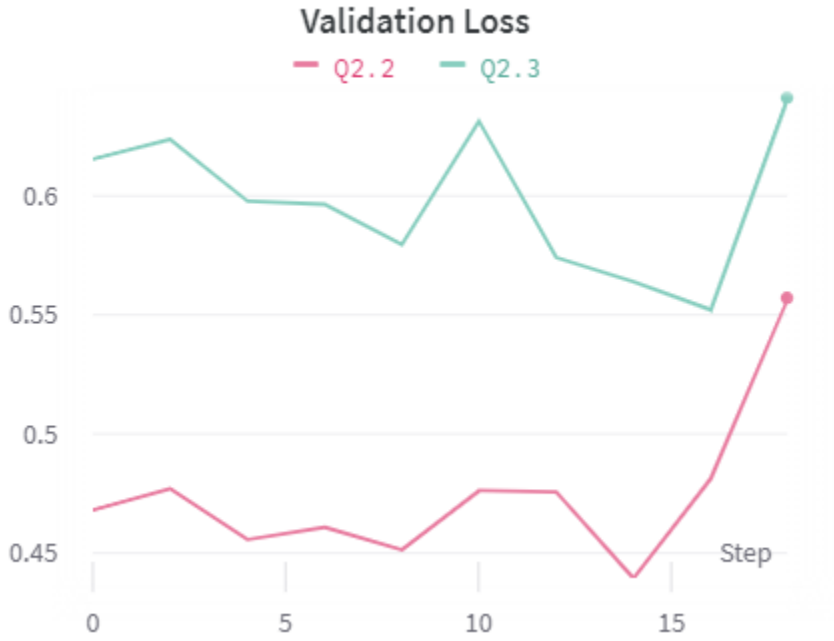


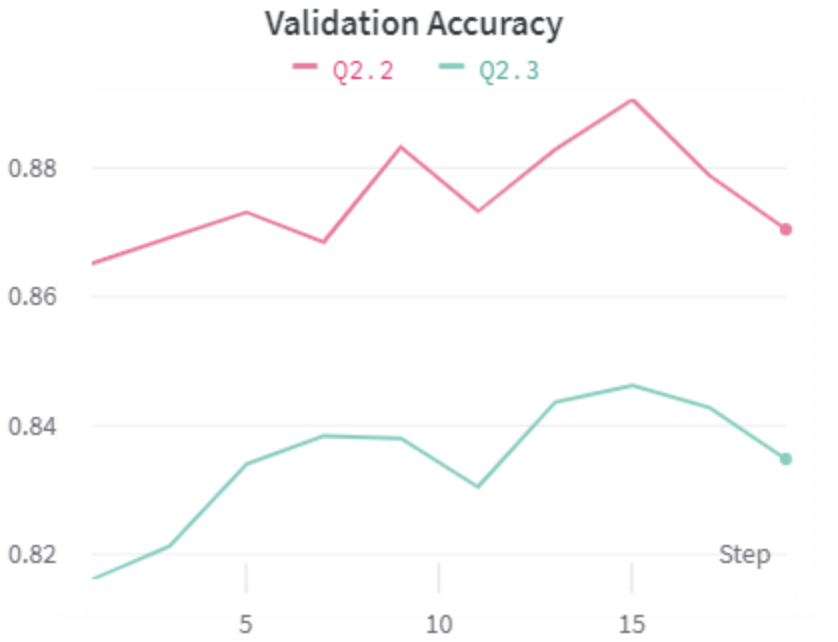
The performance of ResNet50 is greater than the Augmented model for the same reason as in the previous question. The model overfits the data, and due to the augmentations, the performance of the second model is poorer than the base ResNet50.

WandB plots:




Training Accuracy
Q2.2
Q2.3
0.95
0.9
0.85
0.8
Step
5
10
15


Validation Loss
Q2.2
Q2.3
0.6
0.55
0.5
0.45
Step
0
5
10
15


Validation Accuracy
Q2.2
Q2.3
0.88
0.86
0.84
0.82
Step
5
10
15