Assignment 3: Function Approximation with Neural Network and Backpropagation

Process to run the code:

```python
learning_rate = 0.001
optimizer = AdamOptimizer(learning_rate=learning_rate)

layers = [
    DenseLayer(input_size=784, output_size=128, optimizer=optimizer),
    BatchNormalization(input_size=128),
    ReLU(),
    Dropout(dropout_rate=0.2),
    DenseLayer(input_size=128, output_size=64, optimizer=optimizer),
    BatchNormalization(input_size=64),
    ReLU(),
    DenseLayer(input_size=64, output_size=10, optimizer=optimizer),
    Softmax()
]
```

The layers of the feed forward network & the learning rate is defined here. To run the code with a different learning rate or optimizer or layers changes can be made here.

```python
x_train, y_train, x_val, y_val, x_test, y_test = load_data()
train_model(model, x_train, y_train, x_val, y_val, epochs=20, batch_size=64)
test_loss, test_accuracy = evaluate_model(model, x_test, y_test)
print(f'Test Loss: {test_loss:.4f}, Test Accuracy: {test_accuracy:.4f}')
```

Running this portion loads data, trains the model & tests the model & reports accuracy. Inside load_data() function, changes have to be made in case of a different dataset.

Model 1 (Xavier Initialization with 3 hidden layers):
20 epochs, batch size = 64
i) Learning Rate = 0.005:

Epoch 1/20, Training Loss: 0.5885, Training Accuracy: 83.6500, Validation Loss: 0.6194, Validation Accuracy: 83.1250, Validation Macro-F1: 0.8312

Epoch 2/20, Training Loss: 0.6212, Training Accuracy: 85.4437, Validation Loss: 0.6692, Validation Accuracy: 84.9750, Validation Macro-F1: 0.8520

Epoch 3/20, Training Loss: 0.6194, Training Accuracy: 84.6437, Validation Loss: 0.6745, Validation Accuracy: 83.9250, Validation Macro-F1: 0.8416

Epoch 4/20, Training Loss: 0.4881, Training Accuracy: 87.0521, Validation Loss: 0.5281, Validation Accuracy: 86.3833, Validation Macro-F1: 0.8653

Epoch 5/20, Training Loss: 0.4568, Training Accuracy: 87.2979, Validation Loss: 0.4858, Validation Accuracy: 86.5000, Validation Macro-F1: 0.8663

Epoch 6/20, Training Loss: 0.4472, Training Accuracy: 87.9688, Validation Loss: 0.4774, Validation Accuracy: 87.1250, Validation Macro-F1: 0.8724
Epoch 7/20, Training Loss: 0.5526, Training Accuracy: 86.1729, Validation Loss: 0.5886, Validation Accuracy: 85.2000, Validation Macro-F1: 0.8556
Epoch 8/20, Training Loss: 0.4974, Training Accuracy: 87.6208, Validation Loss: 0.5493, Validation Accuracy: 86.5500, Validation Macro-F1: 0.8656
Epoch 9/20, Training Loss: 0.5280, Training Accuracy: 88.5292, Validation Loss: 0.5921, Validation Accuracy: 87.5083, Validation Macro-F1: 0.8755
Epoch 10/20, Training Loss: 0.5484, Training Accuracy: 87.4375, Validation Loss: 0.6027, Validation Accuracy: 86.2833, Validation Macro-F1: 0.8664
Epoch 11/20, Training Loss: 0.5825, Training Accuracy: 88.6396, Validation Loss: 0.6580, Validation Accuracy: 87.5583, Validation Macro-F1: 0.8758
Epoch 12/20, Training Loss: 0.7452, Training Accuracy: 87.7313, Validation Loss: 0.8289, Validation Accuracy: 86.4833, Validation Macro-F1: 0.8667
Epoch 13/20, Training Loss: 0.6677, Training Accuracy: 88.5917, Validation Loss: 0.7609, Validation Accuracy: 87.2917, Validation Macro-F1: 0.8746
Epoch 14/20, Training Loss: 0.6598, Training Accuracy: 89.3854, Validation Loss: 0.7526, Validation Accuracy: 87.9583, Validation Macro-F1: 0.8807
Epoch 15/20, Training Loss: 0.6984, Training Accuracy: 89.4688, Validation Loss: 0.8191, Validation Accuracy: 87.9667, Validation Macro-F1: 0.8809
Epoch 16/20, Training Loss: 0.7324, Training Accuracy: 88.6813, Validation Loss: 0.8492, Validation Accuracy: 87.2250, Validation Macro-F1: 0.8739
Epoch 17/20, Training Loss: 0.8530, Training Accuracy: 88.4958, Validation Loss: 0.9833, Validation Accuracy: 86.9750, Validation Macro-F1: 0.8709
Epoch 18/20, Training Loss: 0.8531, Training Accuracy: 87.5292, Validation Loss: 0.9975, Validation Accuracy: 86.0167, Validation Macro-F1: 0.8623
Epoch 19/20, Training Loss: 0.8434, Training Accuracy: 89.1458, Validation Loss: 0.9763, Validation Accuracy: 87.7750, Validation Macro-F1: 0.8788
Epoch 20/20, Training Loss: 0.9517, Training Accuracy: 88.9312, Validation Loss: 1.1264, Validation Accuracy: 87.2750, Validation Macro-F1: 0.8745
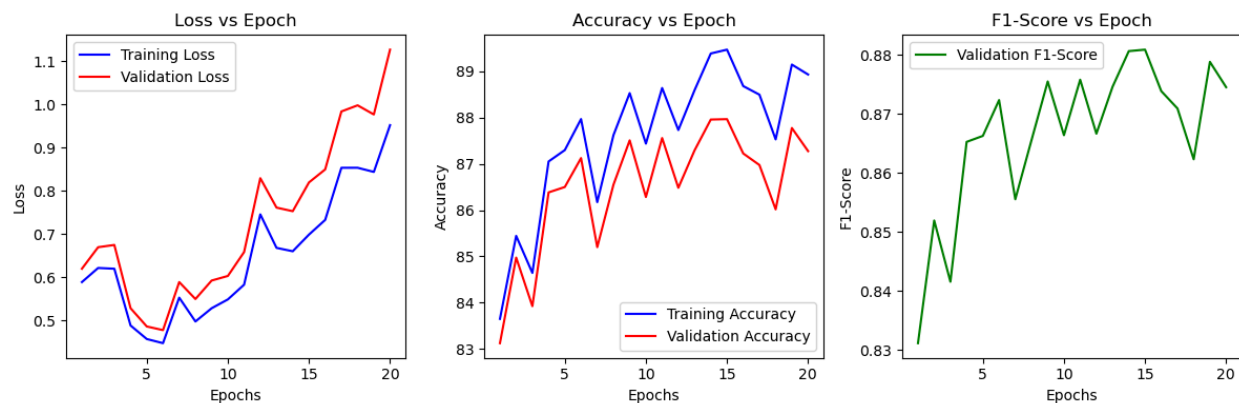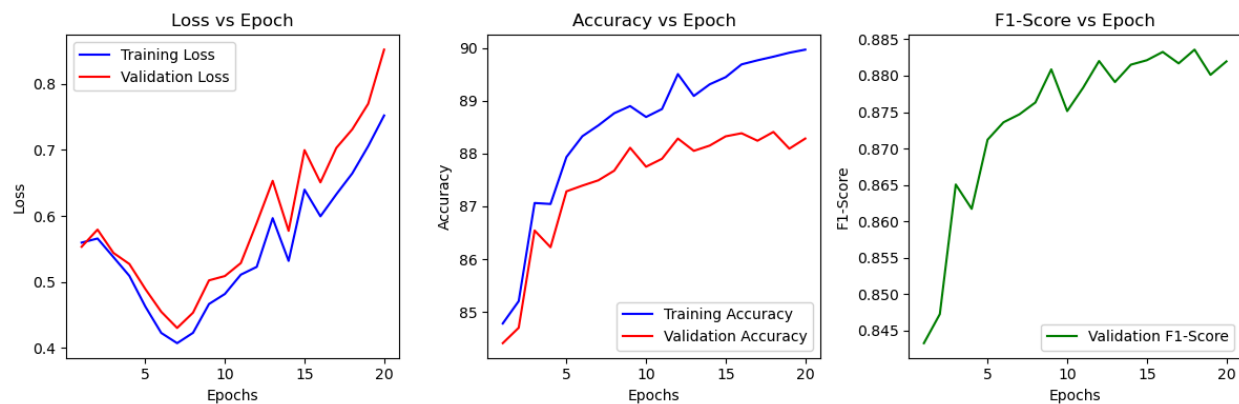

Confusion Matrix for the Best F1-Validation Score Epoch:
[[1021   0   26   32   3   0 116   0   10   0]
 [0 1174   2   17   3   0   7   0   0   0]
 [16   0 920   7 124   0 102   0   6   0]
 [20   3   4 1093   47   1   51   0   4   0]
 [1   3   70   40 990   1 107   0   4   0]
 [1   0   0   2   0 1088   0   31   4   19]
 [154   0   94   12   65   0 897   0   27   0]
 [0   0   0   1   0   49   0 1073   4   36]
 [6   0   4   5   7   2   14   5 1166   2]
 [2   0   0   1   0   26   0   43   1 1134]]

ii) Learning Rate = 0.00375:

Epoch 1/20, Training Loss: 0.5597, Training Accuracy: 84.7812, Validation Loss: 0.5534, Validation Accuracy: 84.4083, Validation Macro-F1: 0.8433
Epoch 2/20, Training Loss: 0.5659, Training Accuracy: 85.2042, Validation Loss: 0.5794, Validation Accuracy: 84.7000, Validation Macro-F1: 0.8473
Epoch 3/20, Training Loss: 0.5377, Training Accuracy: 87.0625, Validation Loss: 0.5441, Validation Accuracy: 86.5417, Validation Macro-F1: 0.8651
Epoch 4/20, Training Loss: 0.5093, Training Accuracy: 87.0438, Validation Loss: 0.5273, Validation Accuracy: 86.2250, Validation Macro-F1: 0.8617
Epoch 5/20, Training Loss: 0.4629, Training Accuracy: 87.9333, Validation Loss: 0.4895, Validation Accuracy: 87.2833, Validation Macro-F1: 0.8712
Epoch 6/20, Training Loss: 0.4229, Training Accuracy: 88.3271, Validation Loss: 0.4550, Validation Accuracy: 87.3917, Validation Macro-F1: 0.8736
Epoch 7/20, Training Loss: 0.4074, Training Accuracy: 88.5312, Validation Loss: 0.4304, Validation Accuracy: 87.4917, Validation Macro-F1: 0.8747
Epoch 8/20, Training Loss: 0.4231, Training Accuracy: 88.7625, Validation Loss: 0.4536, Validation Accuracy: 87.6750, Validation Macro-F1: 0.8763
Epoch 9/20, Training Loss: 0.4667, Training Accuracy: 88.8979, Validation Loss: 0.5025, Validation Accuracy: 88.1083, Validation Macro-F1: 0.8809

Epoch 10/20, Training Loss: 0.4819, Training Accuracy: 88.6917, Validation Loss: 0.5091, Validation Accuracy: 87.7500, Validation Macro-F1: 0.8751
Epoch 11/20, Training Loss: 0.5111, Training Accuracy: 88.8417, Validation Loss: 0.5287, Validation Accuracy: 87.9000, Validation Macro-F1: 0.8784
Epoch 12/20, Training Loss: 0.5230, Training Accuracy: 89.5042, Validation Loss: 0.5895, Validation Accuracy: 88.2833, Validation Macro-F1: 0.8820
Epoch 13/20, Training Loss: 0.5967, Training Accuracy: 89.0896, Validation Loss: 0.6533, Validation Accuracy: 88.0500, Validation Macro-F1: 0.8791
Epoch 14/20, Training Loss: 0.5320, Training Accuracy: 89.3104, Validation Loss: 0.5776, Validation Accuracy: 88.1500, Validation Macro-F1: 0.8815
Epoch 15/20, Training Loss: 0.6399, Training Accuracy: 89.4458, Validation Loss: 0.6998, Validation Accuracy: 88.3250, Validation Macro-F1: 0.8821
Epoch 16/20, Training Loss: 0.5995, Training Accuracy: 89.6875, Validation Loss: 0.6510, Validation Accuracy: 88.3833, Validation Macro-F1: 0.8833
Epoch 17/20, Training Loss: 0.6330, Training Accuracy: 89.7646, Validation Loss: 0.7032, Validation Accuracy: 88.2417, Validation Macro-F1: 0.8817
Epoch 18/20, Training Loss: 0.6646, Training Accuracy: 89.8333, Validation Loss: 0.7314, Validation Accuracy: 88.4083, Validation Macro-F1: 0.8836
Epoch 19/20, Training Loss: 0.7059, Training Accuracy: 89.9083, Validation Loss: 0.7700, Validation Accuracy: 88.0917, Validation Macro-F1: 0.8801
Epoch 20/20, Training Loss: 0.7523, Training Accuracy: 89.9667, Validation Loss: 0.8520, Validation Accuracy: 88.2833, Validation Macro-F1: 0.8820
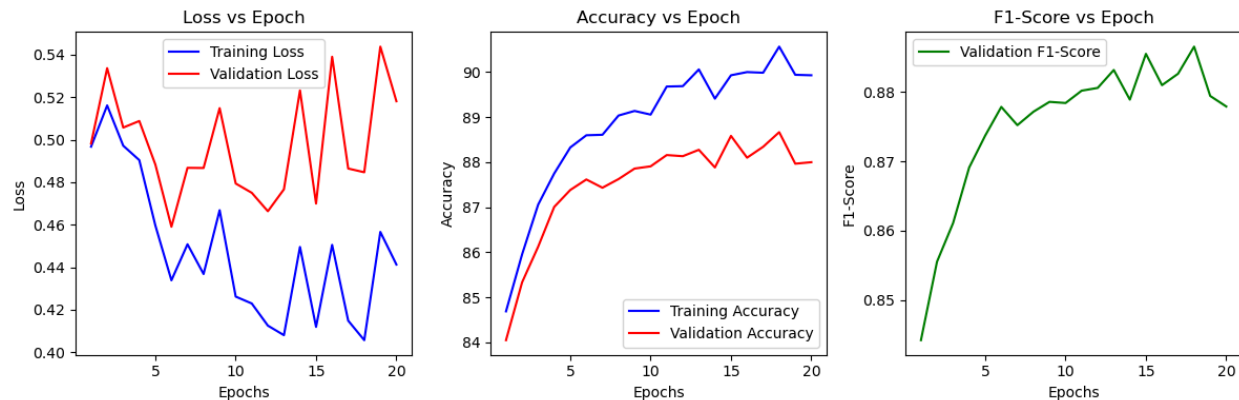

Confusion Matrix for the Best F1-Validation Score Epoch:
```
[[1025    2   14   26    4    1  105    0    8    0]
 [   4 1196    3   13    2    0    3    0    0    0]
 [  17    3  953    7  120    0   91    0    2    0]
 [  44   11    2 1110   58    0   38    0    4    0]
 [   4    2   95   30 1039    3   47    0    1    0]
 [   0    0    0    0    0 1145    0   33    4   12]
 [ 150    2   84   15  101    0  800    0   11    0]
 [   0    0    0    0    0   50    0 1093    2   31]
 [   2    1    9    4   14    4   17    5 1140    2]
 [   1    0    0    0    0   31    0   40    2 1108]]
```

iii) Learning Rate = 0.0025:

Epoch 1/20, Training Loss: 0.4967, Training Accuracy: 84.6896, Validation Loss: 0.4982, Validation Accuracy: 84.0500, Validation Macro-F1: 0.8442
Epoch 2/20, Training Loss: 0.5161, Training Accuracy: 85.9542, Validation Loss: 0.5336, Validation Accuracy: 85.3333, Validation Macro-F1: 0.8555
Epoch 3/20, Training Loss: 0.4972, Training Accuracy: 87.0583, Validation Loss: 0.5057, Validation Accuracy: 86.1250, Validation Macro-F1: 0.8611
Epoch 4/20, Training Loss: 0.4904, Training Accuracy: 87.7500, Validation Loss: 0.5088, Validation Accuracy: 87.0083, Validation Macro-F1: 0.8691
Epoch 5/20, Training Loss: 0.4597, Training Accuracy: 88.3271, Validation Loss: 0.4883, Validation Accuracy: 87.3833, Validation Macro-F1: 0.8737
Epoch 6/20, Training Loss: 0.4339, Training Accuracy: 88.5979, Validation Loss: 0.4591, Validation Accuracy: 87.6167, Validation Macro-F1: 0.8778
Epoch 7/20, Training Loss: 0.4508, Training Accuracy: 88.6083, Validation Loss: 0.4868, Validation Accuracy: 87.4333, Validation Macro-F1: 0.8752
Epoch 8/20, Training Loss: 0.4368, Training Accuracy: 89.0354, Validation Loss: 0.4867, Validation Accuracy: 87.6250, Validation Macro-F1: 0.8771
Epoch 9/20, Training Loss: 0.4668, Training Accuracy: 89.1375, Validation Loss: 0.5148, Validation Accuracy: 87.8583, Validation Macro-F1: 0.8786
Epoch 10/20, Training Loss: 0.4262, Training Accuracy: 89.0563, Validation Loss: 0.4794, Validation Accuracy: 87.9083, Validation Macro-F1: 0.8784
Epoch 11/20, Training Loss: 0.4230, Training Accuracy: 89.6792, Validation Loss: 0.4750, Validation Accuracy: 88.1583, Validation Macro-F1: 0.8802
Epoch 12/20, Training Loss: 0.4126, Training Accuracy: 89.6875, Validation Loss: 0.4664, Validation Accuracy: 88.1333, Validation Macro-F1: 0.8806
Epoch 13/20, Training Loss: 0.4081, Training Accuracy: 90.0583, Validation Loss: 0.4767, Validation Accuracy: 88.2750, Validation Macro-F1: 0.8832

Epoch 14/20, Training Loss: 0.4496, Training Accuracy: 89.4104, Validation Loss: 0.5232, Validation Accuracy: 87.8833, Validation Macro-F1: 0.8789
Epoch 15/20, Training Loss: 0.4120, Training Accuracy: 89.9292, Validation Loss: 0.4700, Validation Accuracy: 88.5833, Validation Macro-F1: 0.8855
Epoch 16/20, Training Loss: 0.4505, Training Accuracy: 90.0000, Validation Loss: 0.5390, Validation Accuracy: 88.1000, Validation Macro-F1: 0.8810
Epoch 17/20, Training Loss: 0.4149, Training Accuracy: 89.9854, Validation Loss: 0.4864, Validation Accuracy: 88.3417, Validation Macro-F1: 0.8826
Epoch 18/20, Training Loss: 0.4057, Training Accuracy: 90.5646, Validation Loss: 0.4847, Validation Accuracy: 88.6667, Validation Macro-F1: 0.8865
Epoch 19/20, Training Loss: 0.4566, Training Accuracy: 89.9396, Validation Loss: 0.5438, Validation Accuracy: 87.9667, Validation Macro-F1: 0.8794
Epoch 20/20, Training Loss: 0.4413, Training Accuracy: 89.9292, Validation Loss: 0.5181, Validation Accuracy: 88.0000, Validation Macro-F1: 0.8779

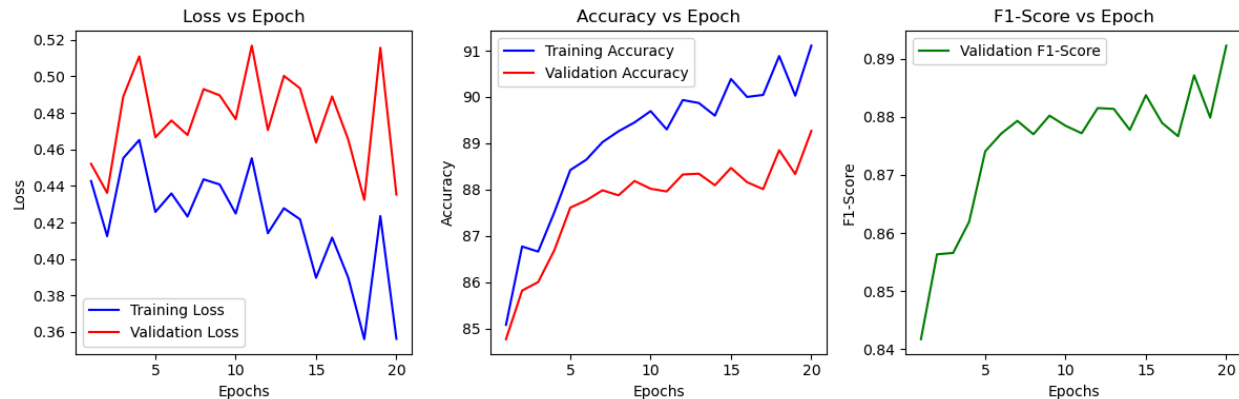Confusion Matrix for the Best F1-Validation Score Epoch:
```
[[ 979    2   21   33    5    0  142    0   18    0]
 [   3 1202    3   28    3    0    4    1    0    0]
 [  18    1 1034   13   81    0   68    0    7    0]
 [  24   11    9 1103   28    0   27    0    5    2]
 [   1    1   99   49 1013    0   48    0    7    0]
 [   0    0    0    1    0 1147    0   32    7   21]
 [ 140    4  117   28   80    0  821    0   14    0]
 [   0    0    0    0    0   30    0 1124    3   22]
 [   1    1    5    2    3    1    8    1 1143    0]
 [   0    0    1    0    0   18    1   56    1 1074]]
```

iv) Learning Rate = 0.001:

Epoch 1/20, Training Loss: 0.4427, Training Accuracy: 85.0771, Validation Loss: 0.4521, Validation Accuracy: 84.7667, Validation Macro-F1: 0.8418
Epoch 2/20, Training Loss: 0.4125, Training Accuracy: 86.7687, Validation Loss: 0.4362, Validation Accuracy: 85.8167, Validation Macro-F1: 0.8563
Epoch 3/20, Training Loss: 0.4552, Training Accuracy: 86.6583, Validation Loss: 0.4889, Validation Accuracy: 86.0000, Validation Macro-F1: 0.8566
Epoch 4/20, Training Loss: 0.4652, Training Accuracy: 87.5021, Validation Loss: 0.5110, Validation Accuracy: 86.6833, Validation Macro-F1: 0.8619
Epoch 5/20, Training Loss: 0.4258, Training Accuracy: 88.4208, Validation Loss: 0.4666, Validation Accuracy: 87.6083, Validation Macro-F1: 0.8741
Epoch 6/20, Training Loss: 0.4359, Training Accuracy: 88.6458, Validation Loss: 0.4758, Validation Accuracy: 87.7667, Validation Macro-F1: 0.8771
Epoch 7/20, Training Loss: 0.4232, Training Accuracy: 89.0229, Validation Loss: 0.4679, Validation Accuracy: 87.9833, Validation Macro-F1: 0.8793
Epoch 8/20, Training Loss: 0.4436, Training Accuracy: 89.2583, Validation Loss: 0.4930, Validation Accuracy: 87.8750, Validation Macro-F1: 0.8770
Epoch 9/20, Training Loss: 0.4408, Training Accuracy: 89.4500, Validation Loss: 0.4896, Validation Accuracy: 88.1833, Validation Macro-F1: 0.8802
Epoch 10/20, Training Loss: 0.4249, Training Accuracy: 89.6938, Validation Loss: 0.4766, Validation Accuracy: 88.0167, Validation Macro-F1: 0.8784
Epoch 11/20, Training Loss: 0.4552, Training Accuracy: 89.2979, Validation Loss: 0.5168, Validation Accuracy: 87.9583, Validation Macro-F1: 0.8772
Epoch 12/20, Training Loss: 0.4141, Training Accuracy: 89.9333, Validation Loss: 0.4706, Validation Accuracy: 88.3250, Validation Macro-F1: 0.8815
Epoch 13/20, Training Loss: 0.4278, Training Accuracy: 89.8687, Validation Loss: 0.5003, Validation Accuracy: 88.3417, Validation Macro-F1: 0.8813
Epoch 14/20, Training Loss: 0.4217, Training Accuracy: 89.5958, Validation Loss: 0.4935, Validation Accuracy: 88.0917, Validation Macro-F1: 0.8777
Epoch 15/20, Training Loss: 0.3897, Training Accuracy: 90.3875, Validation Loss: 0.4638, Validation Accuracy: 88.4667, Validation Macro-F1: 0.8837
Epoch 16/20, Training Loss: 0.4117, Training Accuracy: 90.0000, Validation Loss: 0.4890, Validation Accuracy: 88.1583, Validation Macro-F1: 0.8789

Epoch 17/20, Training Loss: 0.3898, Training Accuracy: 90.0438, Validation Loss: 0.4655, Validation Accuracy: 88.0083, Validation Macro-F1: 0.8767
Epoch 18/20, Training Loss: 0.3561, Training Accuracy: 90.8833, Validation Loss: 0.4323, Validation Accuracy: 88.8500, Validation Macro-F1: 0.8871
Epoch 19/20, Training Loss: 0.4235, Training Accuracy: 90.0271, Validation Loss: 0.5157, Validation Accuracy: 88.3333, Validation Macro-F1: 0.8798
Epoch 20/20, Training Loss: 0.3563, Training Accuracy: 91.1083, Validation Loss: 0.4353, Validation Accuracy: 89.2667, Validation Macro-F1: 0.8922

Confusion Matrix for the Best F1-Validation Score Epoch:

```
[[1046    4   19   28    5    1   82    0   11    2]
 [   3 1168    2   16    0    0    7    0    0    0]
 [  13    1  986   13  116    0   64    0    2    0]
 [  28    5    8 1091   46    0   22    0    0    0]
 [   2    2   96   22 1042    0   44    0    4    0]
 [   0    1    1    1    0 1121    0   25    3   12]
 [ 153    5   89   28   92    0  818    0   17    0]
 [   0    0    0    0    0   28    0 1112    2   38]
 [  12    0    4    8   19    7   10    3 1174    4]
 [   0    0    0    1    0   14    0   43    0 1154]]
```

Model 2 (Kaiming Initialization with 3 hidden layers):

20 epochs, batch size = 64

i) Learning Rate = 0.005:

Epoch 1/20, Training Loss: 0.6432, Training Accuracy: 83.7917, Validation Loss: 0.6456, Validation Accuracy: 83.5500, Validation Macro-F1: 0.8350
Epoch 2/20, Training Loss: 0.6077, Training Accuracy: 86.2750, Validation Loss: 0.5976, Validation Accuracy: 85.9750, Validation Macro-F1: 0.8594
Epoch 3/20, Training Loss: 0.6254, Training Accuracy: 85.8750, Validation Loss: 0.6345, Validation Accuracy: 85.1917, Validation Macro-F1: 0.8506
Epoch 4/20, Training Loss: 0.4823, Training Accuracy: 87.6271, Validation Loss: 0.5025, Validation Accuracy: 86.9917, Validation Macro-F1: 0.8700
Epoch 5/20, Training Loss: 0.4844, Training Accuracy: 87.6687, Validation Loss: 0.5061, Validation Accuracy: 86.7250, Validation Macro-F1: 0.8670
Epoch 6/20, Training Loss: 0.5120, Training Accuracy: 87.7833, Validation Loss: 0.5351, Validation Accuracy: 86.5750, Validation Macro-F1: 0.8646
Epoch 7/20, Training Loss: 0.4921, Training Accuracy: 88.2521, Validation Loss: 0.5266, Validation Accuracy: 86.7917, Validation Macro-F1: 0.8662
Epoch 8/20, Training Loss: 0.5712, Training Accuracy: 87.9375, Validation Loss: 0.6056, Validation Accuracy: 86.7333, Validation Macro-F1: 0.8651
Epoch 9/20, Training Loss: 0.5707, Training Accuracy: 87.8771, Validation Loss: 0.6134, Validation Accuracy: 86.6083, Validation Macro-F1: 0.8636
Epoch 10/20, Training Loss: 0.6285, Training Accuracy: 88.2375, Validation Loss: 0.6694, Validation Accuracy: 87.1583, Validation Macro-F1: 0.8711
Epoch 11/20, Training Loss: 0.7609, Training Accuracy: 88.7312, Validation Loss: 0.8057, Validation Accuracy: 87.6250, Validation Macro-F1: 0.8759
Epoch 12/20, Training Loss: 0.7430, Training Accuracy: 88.6479, Validation Loss: 0.7910, Validation Accuracy: 87.1750, Validation Macro-F1: 0.8709
Epoch 13/20, Training Loss: 0.8167, Training Accuracy: 88.3125, Validation Loss: 0.8852, Validation Accuracy: 86.8667, Validation Macro-F1: 0.8684
Epoch 14/20, Training Loss: 0.7868, Training Accuracy: 89.2729, Validation Loss: 0.8452, Validation Accuracy: 87.8833, Validation Macro-F1: 0.8786
Epoch 15/20, Training Loss: 0.8866, Training Accuracy: 88.7354, Validation Loss: 0.9728, Validation Accuracy: 87.4250, Validation Macro-F1: 0.8757
Epoch 16/20, Training Loss: 0.9885, Training Accuracy: 89.2375, Validation Loss: 1.0481, Validation Accuracy: 88.0167, Validation Macro-F1: 0.8797

Epoch 17/20, Training Loss: 1.0965, Training Accuracy: 88.1854, Validation Loss: 1.2041, Validation Accuracy: 86.8333, Validation Macro-F1: 0.8697
Epoch 18/20, Training Loss: 1.1018, Training Accuracy: 89.3146, Validation Loss: 1.2276, Validation Accuracy: 87.9417, Validation Macro-F1: 0.8791
Epoch 19/20, Training Loss: 1.1630, Training Accuracy: 88.6375, Validation Loss: 1.3018, Validation Accuracy: 87.2583, Validation Macro-F1: 0.8738
Epoch 20/20, Training Loss: 1.0308, Training Accuracy: 88.7396, Validation Loss: 1.1647, Validation Accuracy: 87.4333, Validation Macro-F1: 0.8764
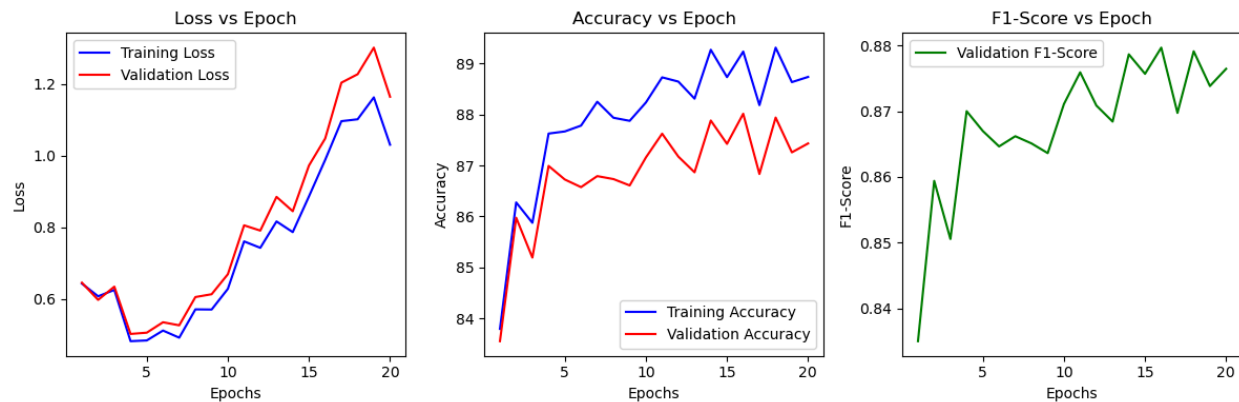
Confusion Matrix for the Best F1-Validation Score Epoch:
[[1035   2   18   43   2   1   87   1   16   0]
 [4 1124   4   41   1   0   6   0   0   0]
 [23   0 951   17 112   1   96   0   6   0]
 [36   4   3 1083   26   1   27   0   5   0]
 [10   2   72   56 951   1   70   0   7   0]
 [0   0   0   0   0 1127   0   51   8   19]
 [187   1 100   24   60   1 822   0   21   0]
 [0   0   0   0   0   19   0 1183   3   28]
 [3   0   3   8   3   2   15   3 1186   1]
 [0   0   0   0   0   23   0   54   0 1100]]

ii) Learning Rate = 0.00375:

Epoch 1/20, Training Loss: 0.5571, Training Accuracy: 83.9979, Validation Loss: 0.5901, Validation Accuracy: 83.2667, Validation Macro-F1: 0.8250
Epoch 2/20, Training Loss: 0.5699, Training Accuracy: 86.4188, Validation Loss: 0.6049, Validation Accuracy: 85.5833, Validation Macro-F1: 0.8557
Epoch 3/20, Training Loss: 0.5424, Training Accuracy: 87.1042, Validation Loss: 0.5912, Validation Accuracy: 86.0167, Validation Macro-F1: 0.8588
Epoch 4/20, Training Loss: 0.4799, Training Accuracy: 88.0042, Validation Loss: 0.5213, Validation Accuracy: 86.8500, Validation Macro-F1: 0.8676
Epoch 5/20, Training Loss: 0.4624, Training Accuracy: 88.3312, Validation Loss: 0.5081, Validation Accuracy: 86.9417, Validation Macro-F1: 0.8700
Epoch 6/20, Training Loss: 0.4613, Training Accuracy: 88.4375, Validation Loss: 0.5112, Validation Accuracy: 86.8417, Validation Macro-F1: 0.8674
Epoch 7/20, Training Loss: 0.4939, Training Accuracy: 87.6813, Validation Loss: 0.5528, Validation Accuracy: 86.4167, Validation Macro-F1: 0.8614
Epoch 8/20, Training Loss: 0.4277, Training Accuracy: 88.9833, Validation Loss: 0.4733, Validation Accuracy: 87.6333, Validation Macro-F1: 0.8769
Epoch 9/20, Training Loss: 0.4747, Training Accuracy: 88.7500, Validation Loss: 0.5441, Validation Accuracy: 86.8667, Validation Macro-F1: 0.8688
Epoch 10/20, Training Loss: 0.4378, Training Accuracy: 89.1437, Validation Loss: 0.5012, Validation Accuracy: 87.6417, Validation Macro-F1: 0.8764
Epoch 11/20, Training Loss: 0.4771, Training Accuracy: 89.1625, Validation Loss: 0.5413, Validation Accuracy: 87.7833, Validation Macro-F1: 0.8776
Epoch 12/20, Training Loss: 0.5257, Training Accuracy: 89.1292, Validation Loss: 0.6020, Validation Accuracy: 87.6583, Validation Macro-F1: 0.8775
Epoch 13/20, Training Loss: 0.4644, Training Accuracy: 89.3729, Validation Loss: 0.5475, Validation Accuracy: 87.7417, Validation Macro-F1: 0.8783
Epoch 14/20, Training Loss: 0.5439, Training Accuracy: 89.5854, Validation Loss: 0.6221, Validation Accuracy: 87.7500, Validation Macro-F1: 0.8776
Epoch 15/20, Training Loss: 0.6221, Training Accuracy: 88.7188, Validation Loss: 0.7197, Validation Accuracy: 86.9250, Validation Macro-F1: 0.8680
Epoch 16/20, Training Loss: 0.5222, Training Accuracy: 89.7625, Validation Loss: 0.6082, Validation Accuracy: 87.9833, Validation Macro-F1: 0.8793
Epoch 17/20, Training Loss: 0.5638, Training Accuracy: 89.7708, Validation Loss: 0.6542, Validation Accuracy: 87.6583, Validation Macro-F1: 0.8770
Epoch 18/20, Training Loss: 0.6369, Training Accuracy: 89.1063, Validation Loss: 0.7580, Validation Accuracy: 87.3583, Validation Macro-F1: 0.8732

Epoch 19/20, Training Loss: 0.6488, Training Accuracy: 89.7292, Validation Loss: 0.7630, Validation Accuracy: 87.8417, Validation Macro-F1: 0.8794
Epoch 20/20, Training Loss: 0.7070, Training Accuracy: 89.7667, Validation Loss: 0.8130, Validation Accuracy: 88.0167, Validation Macro-F1: 0.8803
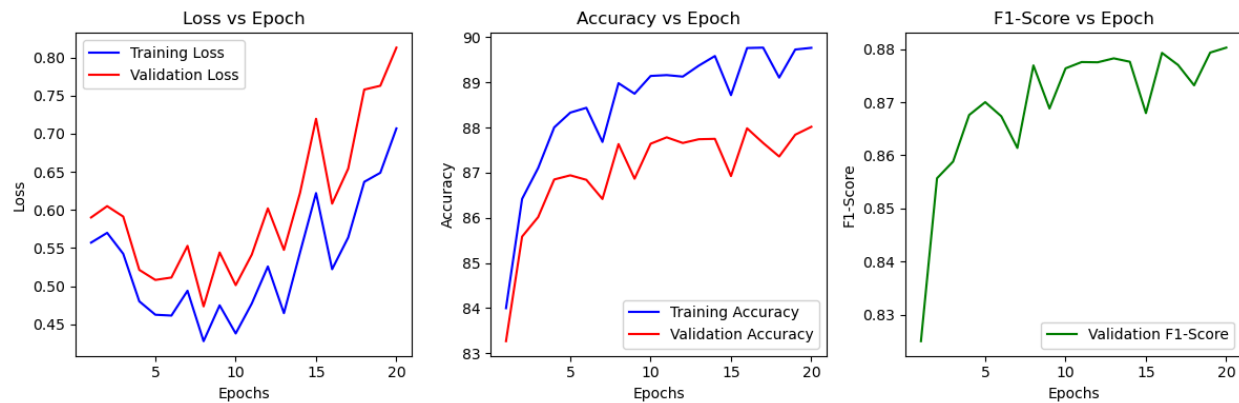
Confusion Matrix for the Best F1-Validation Score Epoch:

```
[[ 996    6   17   38    0    0  144    1   15    0]
 [   3 1129    1   13    1    1    5    0    1    0]
 [  20    3 1030   10   45    1   79    0    8    0]
 [  31   13    4 1109   42    0   41    0    2    0]
 [   2    4  162   25  871    0   83    0    8    0]
 [   0    1    0    1    1 1170    1   40    4   33]
 [ 150    2  104   22   69    0  861    0    8    1]
 [   0    0    0    0    0   12    0 1138    3   53]
 [   2    0    9    4    6    6   17    7 1170    2]
 [   0    1    0    0    1   12    0   36    1 1088]]
```

iii) Learning Rate = 0.0025:

Epoch 1/20, Training Loss: 0.4936, Training Accuracy: 85.0958, Validation Loss: 0.5172, Validation Accuracy: 84.5667, Validation Macro-F1: 0.8408
Epoch 2/20, Training Loss: 0.4893, Training Accuracy: 85.7771, Validation Loss: 0.5318, Validation Accuracy: 84.9000, Validation Macro-F1: 0.8489
Epoch 3/20, Training Loss: 0.4723, Training Accuracy: 87.4646, Validation Loss: 0.5211, Validation Accuracy: 86.1250, Validation Macro-F1: 0.8614
Epoch 4/20, Training Loss: 0.4766, Training Accuracy: 87.8750, Validation Loss: 0.5344, Validation Accuracy: 86.4250, Validation Macro-F1: 0.8623
Epoch 5/20, Training Loss: 0.4528, Training Accuracy: 88.3187, Validation Loss: 0.5173, Validation Accuracy: 86.7583, Validation Macro-F1: 0.8672
Epoch 6/20, Training Loss: 0.4584, Training Accuracy: 88.5271, Validation Loss: 0.5304, Validation Accuracy: 87.0333, Validation Macro-F1: 0.8688
Epoch 7/20, Training Loss: 0.4497, Training Accuracy: 88.8167, Validation Loss: 0.5329, Validation Accuracy: 87.1500, Validation Macro-F1: 0.8704
Epoch 8/20, Training Loss: 0.4227, Training Accuracy: 88.4854, Validation Loss: 0.4910, Validation Accuracy: 87.0000, Validation Macro-F1: 0.8700
Epoch 9/20, Training Loss: 0.4668, Training Accuracy: 88.6354, Validation Loss: 0.5581, Validation Accuracy: 87.0583, Validation Macro-F1: 0.8685
Epoch 10/20, Training Loss: 0.4426, Training Accuracy: 89.3792, Validation Loss: 0.5400, Validation Accuracy: 87.4583, Validation Macro-F1: 0.8728
Epoch 11/20, Training Loss: 0.4116, Training Accuracy: 89.5062, Validation Loss: 0.5051, Validation Accuracy: 87.6250, Validation Macro-F1: 0.8740
Epoch 12/20, Training Loss: 0.4030, Training Accuracy: 89.8063, Validation Loss: 0.4934, Validation Accuracy: 87.8083, Validation Macro-F1: 0.8776
Epoch 13/20, Training Loss: 0.4307, Training Accuracy: 89.9000, Validation Loss: 0.5175, Validation Accuracy: 87.7583, Validation Macro-F1: 0.8766
Epoch 14/20, Training Loss: 0.4231, Training Accuracy: 90.1167, Validation Loss: 0.5262, Validation Accuracy: 87.9167, Validation Macro-F1: 0.8778
Epoch 15/20, Training Loss: 0.4556, Training Accuracy: 89.9188, Validation Loss: 0.5715, Validation Accuracy: 87.7833, Validation Macro-F1: 0.8760
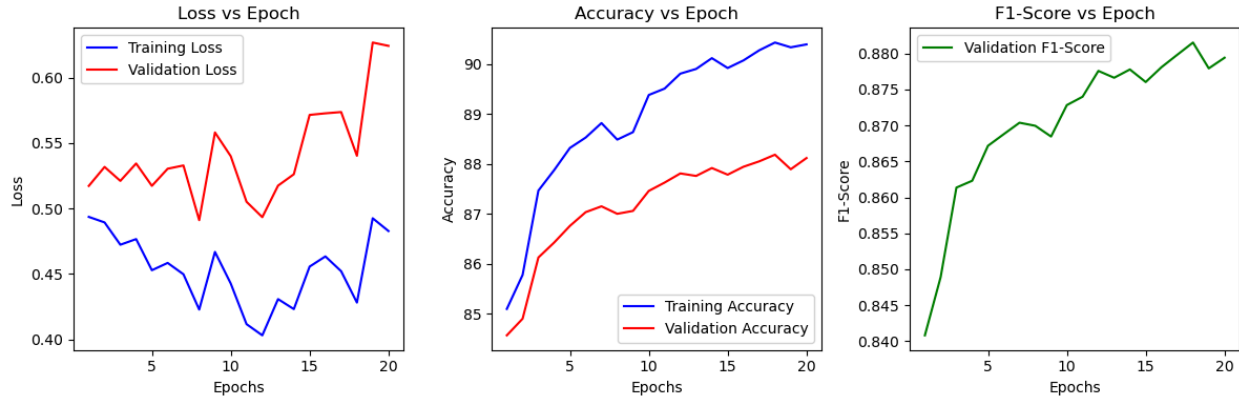Epoch 16/20, Training Loss: 0.4633, Training Accuracy: 90.0708, Validation Loss: 0.5727, Validation Accuracy: 87.9417, Validation Macro-F1: 0.8781
Epoch 17/20, Training Loss: 0.4520, Training Accuracy: 90.2729, Validation Loss: 0.5737, Validation Accuracy: 88.0500, Validation Macro-F1: 0.8798
Epoch 18/20, Training Loss: 0.4280, Training Accuracy: 90.4292, Validation Loss: 0.5403, Validation Accuracy: 88.1833, Validation Macro-F1: 0.8815
Epoch 19/20, Training Loss: 0.4925, Training Accuracy: 90.3333, Validation Loss: 0.6269, Validation Accuracy: 87.8917, Validation Macro-F1: 0.8779
Epoch 20/20, Training Loss: 0.4828, Training Accuracy: 90.3917, Validation Loss: 0.6243, Validation Accuracy: 88.1167, Validation Macro-F1: 0.8794
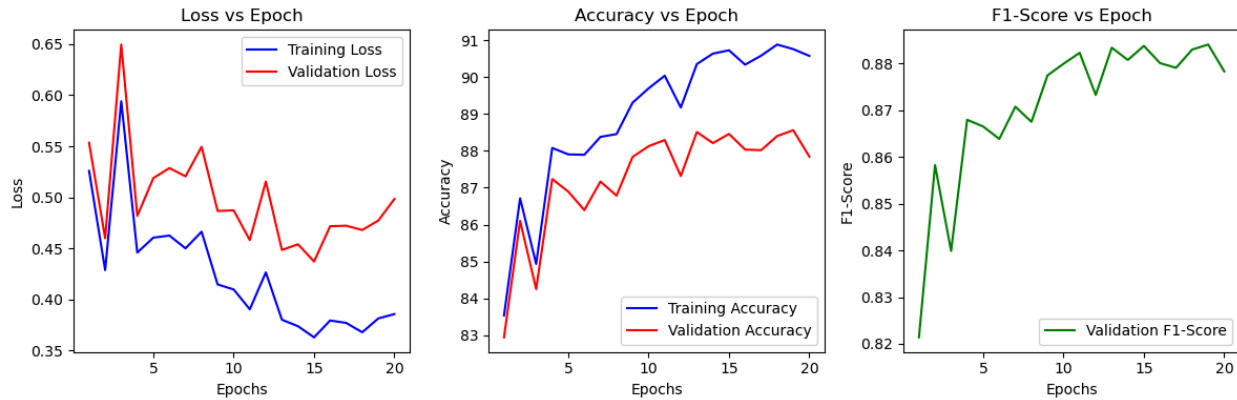
Confusion Matrix for the Best F1-Validation Score Epoch:
```
[[1002    0   12   30    5    1  112    0   16    2]
 [   4 1102    2   18    2    1    2    0    0    0]
 [  15    3 1011   15   90    2   75    0    6    0]
 [  40    6   12 1094   27    0   30    0    0    1]
 [   0    0  131   42  910    1   82    0    4    0]
 [   0    0    1    1    0 1106    0   55    5   19]
 [ 146    3  136   30   57    0  824    0   15    0]
 [   0    0    0    0    0   23    0 1141    2   36]
 [   6    1    8    9    6    4   10    4 1195    0]
 [   0    0    0    0    0    6    1   44    1 1197]]
```

iv) Learning Rate = 0.001:

Epoch 1/20, Training Loss: 0.5258, Training Accuracy: 83.5354, Validation Loss: 0.5535, Validation Accuracy: 82.9417, Validation Macro-F1: 0.8214
Epoch 2/20, Training Loss: 0.4288, Training Accuracy: 86.7104, Validation Loss: 0.4600, Validation Accuracy: 86.1000, Validation Macro-F1: 0.8583
Epoch 3/20, Training Loss: 0.5940, Training Accuracy: 84.9313, Validation Loss: 0.6496, Validation Accuracy: 84.2500, Validation Macro-F1: 0.8399
Epoch 4/20, Training Loss: 0.4459, Training Accuracy: 88.0771, Validation Loss: 0.4819, Validation Accuracy: 87.2333, Validation Macro-F1: 0.8680
Epoch 5/20, Training Loss: 0.4606, Training Accuracy: 87.9021, Validation Loss: 0.5189, Validation Accuracy: 86.8917, Validation Macro-F1: 0.8665
Epoch 6/20, Training Loss: 0.4626, Training Accuracy: 87.8917, Validation Loss: 0.5286, Validation Accuracy: 86.3917, Validation Macro-F1: 0.8639
Epoch 7/20, Training Loss: 0.4501, Training Accuracy: 88.3771, Validation Loss: 0.5205, Validation Accuracy: 87.1667, Validation Macro-F1: 0.8708
Epoch 8/20, Training Loss: 0.4664, Training Accuracy: 88.4521, Validation Loss: 0.5495, Validation Accuracy: 86.7833, Validation Macro-F1: 0.8675
Epoch 9/20, Training Loss: 0.4147, Training Accuracy: 89.3042, Validation Loss: 0.4867, Validation Accuracy: 87.8333, Validation Macro-F1: 0.8775
Epoch 10/20, Training Loss: 0.4098, Training Accuracy: 89.6938, Validation Loss: 0.4872, Validation Accuracy: 88.1250, Validation Macro-F1: 0.8799
Epoch 11/20, Training Loss: 0.3904, Training Accuracy: 90.0354, Validation Loss: 0.4581, Validation Accuracy: 88.2917, Validation Macro-F1: 0.8823
Epoch 12/20, Training Loss: 0.4265, Training Accuracy: 89.1729, Validation Loss: 0.5155, Validation Accuracy: 87.3167, Validation Macro-F1: 0.8733
Epoch 13/20, Training Loss: 0.3802, Training Accuracy: 90.3521, Validation Loss: 0.4485, Validation Accuracy: 88.5083, Validation Macro-F1: 0.8834
Epoch 14/20, Training Loss: 0.3738, Training Accuracy: 90.6354, Validation Loss: 0.4541, Validation Accuracy: 88.2083, Validation Macro-F1: 0.8808
Epoch 15/20, Training Loss: 0.3630, Training Accuracy: 90.7250, Validation Loss: 0.4372, Validation Accuracy: 88.4583, Validation Macro-F1: 0.8838
Epoch 16/20, Training Loss: 0.3795, Training Accuracy: 90.3354, Validation Loss: 0.4717, Validation Accuracy: 88.0333, Validation Macro-F1: 0.8801
Epoch 17/20, Training Loss: 0.3770, Training Accuracy: 90.5792, Validation Loss: 0.4722, Validation Accuracy: 88.0167, Validation Macro-F1: 0.8791
Epoch 18/20, Training Loss: 0.3679, Training Accuracy: 90.8813, Validation Loss: 0.4680, Validation Accuracy: 88.4000, Validation Macro-F1: 0.8830
Epoch 19/20, Training Loss: 0.3815, Training Accuracy: 90.7563, Validation Loss: 0.4773, Validation Accuracy: 88.5583, Validation Macro-F1: 0.8841
Epoch 20/20, Training Loss: 0.3857, Training Accuracy: 90.5729, Validation Loss: 0.4984, Validation Accuracy: 87.8417, Validation Macro-F1: 0.8783

Confusion Matrix for the Best F1-Validation Score Epoch:

[[1022   2   21   17   2   0   98   0   16   0]
 [4 1139   2   23   5   0   5   0   0   1]
 [20   0 1021   4 139   1   38   0   14   0]
 [42   8   13 1036   84   0   30   0   2   1]
 [1   2   84   13 989   1   43   0   5   0]
 [0   1   0   0   0 1201   0   30   4   8]
 [163   2 100   19 113   0 758   0   25   0]
 [0   0   0   0   0   22   0 1130   1   49]
 [4   0   7   4   5   1   15   5 1151   2]
 [0   0   0   0   0   14   0   38   0 1180]]

Model 3 (Xavier Initialization with 5 hidden layers):

20 epochs, batch size = 64

i) Learning Rate = 0.005:

Epoch 1/20, Training Loss: 0.5858, Training Accuracy: 83.1792, Validation Loss: 0.5836, Validation Accuracy: 83.2167, Validation Macro-F1: 0.8329
Epoch 2/20, Training Loss: 0.5310, Training Accuracy: 84.7521, Validation Loss: 0.5460, Validation Accuracy: 84.5167, Validation Macro-F1: 0.8489
Epoch 3/20, Training Loss: 0.5052, Training Accuracy: 86.2604, Validation Loss: 0.5297, Validation Accuracy: 85.7917, Validation Macro-F1: 0.8593
Epoch 4/20, Training Loss: 0.4918, Training Accuracy: 86.5208, Validation Loss: 0.5164, Validation Accuracy: 85.9500, Validation Macro-F1: 0.8602
Epoch 5/20, Training Loss: 0.4925, Training Accuracy: 86.4583, Validation Loss: 0.5281, Validation Accuracy: 85.6750, Validation Macro-F1: 0.8597
Epoch 6/20, Training Loss: 0.4830, Training Accuracy: 87.3917, Validation Loss: 0.5260, Validation Accuracy: 86.9417, Validation Macro-F1: 0.8697
Epoch 7/20, Training Loss: 0.5214, Training Accuracy: 87.8083, Validation Loss: 0.5880, Validation Accuracy: 86.5833, Validation Macro-F1: 0.8675
Epoch 8/20, Training Loss: 0.5222, Training Accuracy: 87.3625, Validation Loss: 0.5741, Validation Accuracy: 86.5833, Validation Macro-F1: 0.8665
Epoch 9/20, Training Loss: 0.5192, Training Accuracy: 88.2271, Validation Loss: 0.5675, Validation Accuracy: 87.1750, Validation Macro-F1: 0.8714
Epoch 10/20, Training Loss: 0.5687, Training Accuracy: 88.4646, Validation Loss: 0.6466, Validation Accuracy: 87.2917, Validation Macro-F1: 0.8735
Epoch 11/20, Training Loss: 0.5548, Training Accuracy: 88.5875, Validation Loss: 0.6327, Validation Accuracy: 87.3333, Validation Macro-F1: 0.8738
Epoch 12/20, Training Loss: 0.6653, Training Accuracy: 88.5958, Validation Loss: 0.7645, Validation Accuracy: 87.3750, Validation Macro-F1: 0.8737
Epoch 13/20, Training Loss: 0.9588, Training Accuracy: 88.1979, Validation Loss: 1.0879, Validation Accuracy: 86.8917, Validation Macro-F1: 0.8704
Epoch 14/20, Training Loss: 0.7366, Training Accuracy: 89.0646, Validation Loss: 0.8621, Validation Accuracy: 87.5000, Validation Macro-F1: 0.8760
Epoch 15/20, Training Loss: 1.0964, Training Accuracy: 88.9292, Validation Loss: 1.2820, Validation Accuracy: 87.3667, Validation Macro-F1: 0.8733
Epoch 16/20, Training Loss: 1.1268, Training Accuracy: 88.8167, Validation Loss: 1.3827, Validation Accuracy: 87.0417, Validation Macro-F1: 0.8719
Epoch 17/20, Training Loss: 0.8495, Training Accuracy: 89.0771, Validation Loss: 0.9789, Validation Accuracy: 87.6083, Validation Macro-F1: 0.8764
Epoch 18/20, Training Loss: 0.8416, Training Accuracy: 89.0708, Validation Loss: 0.9718, Validation Accuracy: 87.3333, Validation Macro-F1: 0.8757
Epoch 19/20, Training Loss: 1.2424, Training Accuracy: 88.7729, Validation Loss: 1.4202, Validation Accuracy: 87.6500, Validation Macro-F1: 0.8759
Epoch 20/20, Training Loss: 0.9423, Training Accuracy: 89.5958, Validation Loss: 1.1188, Validation Accuracy: 88.0333, Validation Macro-F1: 0.8806
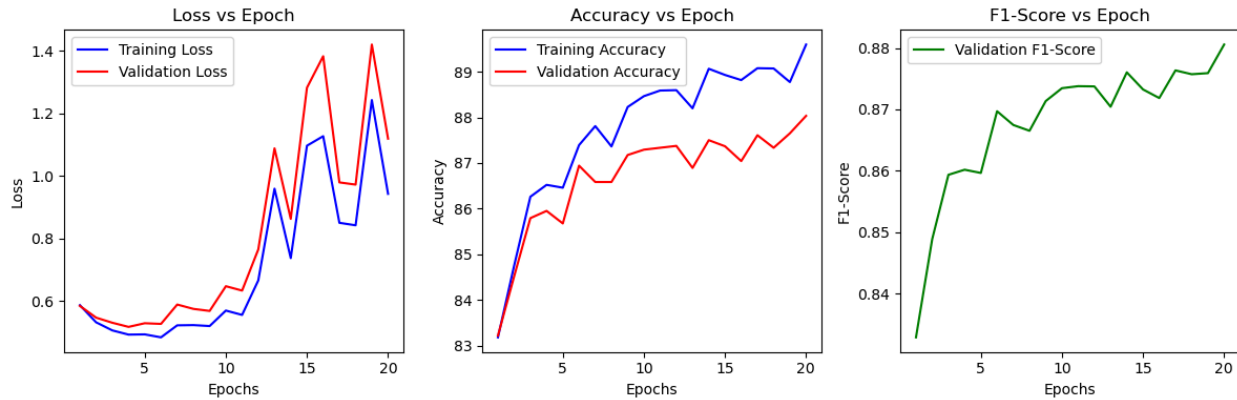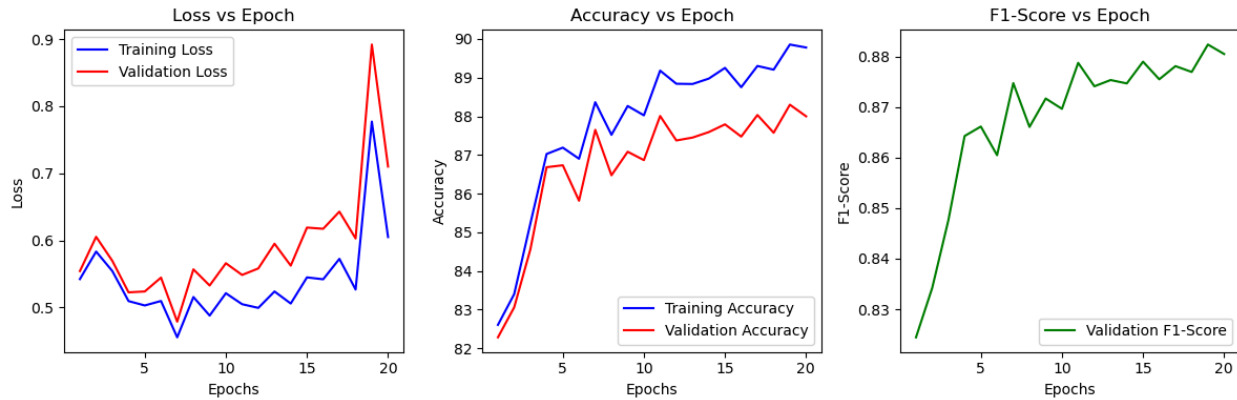
Confusion Matrix for the Best F1-Validation Score Epoch:
```
[[1010    0   19   57    1    1  102    0   10    0]
 [   3 1127    6   32    3    0    0    0    0    0]
 [  11    1 1048   16   82    1   60    0    2    0]
 [  36    6   11 1097   38    1   14    0    3    0]
 [   1    0  123   35  995    1   53    0   10    0]
 [   0    0    1    2    1 1111    0   46    6   30]
 [ 152    1  127   38   85    0  805    0   13    0]
 [   0    0    0    0    0   31    0 1132    2   48]
 [   5    1   11    8    5    3   11    6 1129    2]
 [   0    0    1    0    0   12    0   49    0 1110]]
```

ii) Learning Rate = 0.00375:

Epoch 1/20, Training Loss: 0.5424, Training Accuracy: 82.6063, Validation Loss: 0.5544, Validation Accuracy: 82.2833, Validation Macro-F1: 0.8244
Epoch 2/20, Training Loss: 0.5834, Training Accuracy: 83.4000, Validation Loss: 0.6054, Validation Accuracy: 83.0583, Validation Macro-F1: 0.8342
Epoch 3/20, Training Loss: 0.5544, Training Accuracy: 85.2271, Validation Loss: 0.5696, Validation Accuracy: 84.5583, Validation Macro-F1: 0.8476
Epoch 4/20, Training Loss: 0.5096, Training Accuracy: 87.0250, Validation Loss: 0.5227, Validation Accuracy: 86.6833, Validation Macro-F1: 0.8643
Epoch 5/20, Training Loss: 0.5032, Training Accuracy: 87.1896, Validation Loss: 0.5240, Validation Accuracy: 86.7333, Validation Macro-F1: 0.8661
Epoch 6/20, Training Loss: 0.5098, Training Accuracy: 86.9021, Validation Loss: 0.5447, Validation Accuracy: 85.8167, Validation Macro-F1: 0.8605
Epoch 7/20, Training Loss: 0.4557, Training Accuracy: 88.3646, Validation Loss: 0.4791, Validation Accuracy: 87.6500, Validation Macro-F1: 0.8747
Epoch 8/20, Training Loss: 0.5157, Training Accuracy: 87.5208, Validation Loss: 0.5570, Validation Accuracy: 86.4750, Validation Macro-F1: 0.8661
Epoch 9/20, Training Loss: 0.4881, Training Accuracy: 88.2667, Validation Loss: 0.5329, Validation Accuracy: 87.0833, Validation Macro-F1: 0.8717
Epoch 10/20, Training Loss: 0.5214, Training Accuracy: 88.0250, Validation Loss: 0.5661, Validation Accuracy: 86.8667, Validation Macro-F1: 0.8696
Epoch 11/20, Training Loss: 0.5051, Training Accuracy: 89.1813, Validation Loss: 0.5486, Validation Accuracy: 88.0083, Validation Macro-F1: 0.8787
Epoch 12/20, Training Loss: 0.4995, Training Accuracy: 88.8417, Validation Loss: 0.5583, Validation Accuracy: 87.3750, Validation Macro-F1: 0.8741
Epoch 13/20, Training Loss: 0.5240, Training Accuracy: 88.8354, Validation Loss: 0.5951, Validation Accuracy: 87.4500, Validation Macro-F1: 0.8753
Epoch 14/20, Training Loss: 0.5060, Training Accuracy: 88.9750, Validation Loss: 0.5625, Validation Accuracy: 87.5917, Validation Macro-F1: 0.8747
Epoch 15/20, Training Loss: 0.5450, Training Accuracy: 89.2521, Validation Loss: 0.6192, Validation Accuracy: 87.7917, Validation Macro-F1: 0.8790
Epoch 16/20, Training Loss: 0.5421, Training Accuracy: 88.7521, Validation Loss: 0.6174, Validation Accuracy: 87.4750, Validation Macro-F1: 0.8755
Epoch 17/20, Training Loss: 0.5726, Training Accuracy: 89.3021, Validation Loss: 0.6429, Validation Accuracy: 88.0333, Validation Macro-F1: 0.8781
Epoch 18/20, Training Loss: 0.5269, Training Accuracy: 89.2062, Validation Loss: 0.6029, Validation Accuracy: 87.5750, Validation Macro-F1: 0.8769
Epoch 19/20, Training Loss: 0.7771, Training Accuracy: 89.8563, Validation Loss: 0.8919, Validation Accuracy: 88.3000, Validation Macro-F1: 0.8824
Epoch 20/20, Training Loss: 0.6051, Training Accuracy: 89.7792, Validation Loss: 0.7101, Validation Accuracy: 88.0000, Validation Macro-F1: 0.8805
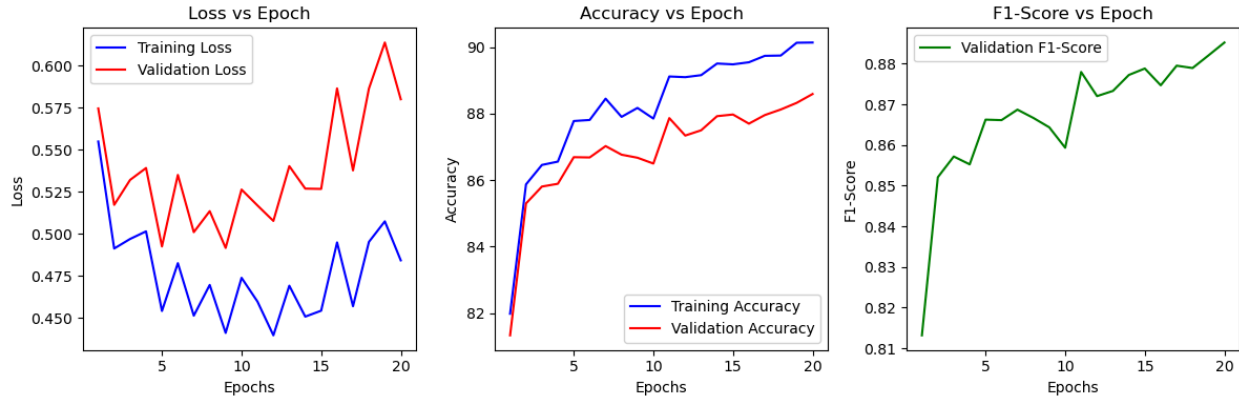
Confusion Matrix for the Best F1-Validation Score Epoch:

```
[[ 998    1   14   46    2    3  113    0    8    2]
 [   3 1188    3   23    4    0    6    0    1    2]
 [  11    0 1006   11   70    0   85    0    3    1]
 [  18    2    5 1154   55    0   35    0    6    0]
 [   2    0  134   32  986    0   71    0    4    0]
 [   1    0    0    1    0 1060    1   45    6   31]
 [ 153    1  113   35   75    0  775    0   11    1]
 [   0    0    0    0    0   22    0 1136    2   32]
 [   4    0    8    4    6    2   20    5 1157    1]
 [   0    0    0    0    0    8    0   40    0 1136]]
```

iii) Learning Rate = 0.0025:

Epoch 1/20, Training Loss: 0.5549, Training Accuracy: 81.9833, Validation Loss: 0.5746, Validation Accuracy: 81.3333, Validation Macro-F1: 0.8132

Epoch 2/20, Training Loss: 0.4913, Training Accuracy: 85.8729, Validation Loss: 0.5173, Validation Accuracy: 85.3000, Validation Macro-F1: 0.8521

Epoch 3/20, Training Loss: 0.4970, Training Accuracy: 86.4583, Validation Loss: 0.5321, Validation Accuracy: 85.8083, Validation Macro-F1: 0.8572

Epoch 4/20, Training Loss: 0.5015, Training Accuracy: 86.5583, Validation Loss: 0.5392, Validation Accuracy: 85.8917, Validation Macro-F1: 0.8552

Epoch 5/20, Training Loss: 0.4541, Training Accuracy: 87.7792, Validation Loss: 0.4925, Validation Accuracy: 86.6917, Validation Macro-F1: 0.8662

Epoch 6/20, Training Loss: 0.4826, Training Accuracy: 87.8104, Validation Loss: 0.5351, Validation Accuracy: 86.6833, Validation Macro-F1: 0.8661

Epoch 7/20, Training Loss: 0.4514, Training Accuracy: 88.4521, Validation Loss: 0.5010, Validation Accuracy: 87.0250, Validation Macro-F1: 0.8687

Epoch 8/20, Training Loss: 0.4697, Training Accuracy: 87.9042, Validation Loss: 0.5135, Validation Accuracy: 86.7667, Validation Macro-F1: 0.8667

Epoch 9/20, Training Loss: 0.4411, Training Accuracy: 88.1729, Validation Loss: 0.4916, Validation Accuracy: 86.6750, Validation Macro-F1: 0.8643

Epoch 10/20, Training Loss: 0.4739, Training Accuracy: 87.8521, Validation Loss: 0.5264, Validation Accuracy: 86.5000, Validation Macro-F1: 0.8593

Epoch 11/20, Training Loss: 0.4597, Training Accuracy: 89.1188, Validation Loss: 0.5169, Validation Accuracy: 87.8667, Validation Macro-F1: 0.8780

Epoch 12/20, Training Loss: 0.4397, Training Accuracy: 89.0979, Validation Loss: 0.5077, Validation Accuracy: 87.3417, Validation Macro-F1: 0.8720

Epoch 13/20, Training Loss: 0.4692, Training Accuracy: 89.1583, Validation Loss: 0.5403, Validation Accuracy: 87.5000, Validation Macro-F1: 0.8733

Epoch 14/20, Training Loss: 0.4508, Training Accuracy: 89.5083, Validation Loss: 0.5269, Validation Accuracy: 87.9250, Validation Macro-F1: 0.8772

Epoch 15/20, Training Loss: 0.4543, Training Accuracy: 89.4833, Validation Loss: 0.5267, Validation Accuracy: 87.9750, Validation Macro-F1: 0.8788

Epoch 16/20, Training Loss: 0.4949, Training Accuracy: 89.5458, Validation Loss: 0.5865, Validation Accuracy: 87.7000, Validation Macro-F1: 0.8747

Epoch 17/20, Training Loss: 0.4570, Training Accuracy: 89.7396, Validation Loss: 0.5377, Validation Accuracy: 87.9583, Validation Macro-F1: 0.8795

Epoch 18/20, Training Loss: 0.4952, Training Accuracy: 89.7479, Validation Loss: 0.5864, Validation Accuracy: 88.1250, Validation Macro-F1: 0.8789

Epoch 19/20, Training Loss: 0.5075, Training Accuracy: 90.1333, Validation Loss: 0.6138, Validation Accuracy: 88.3250, Validation Macro-F1: 0.8820

Epoch 20/20, Training Loss: 0.4843, Training Accuracy: 90.1396, Validation Loss: 0.5801, Validation Accuracy: 88.5917, Validation Macro-F1: 0.8852
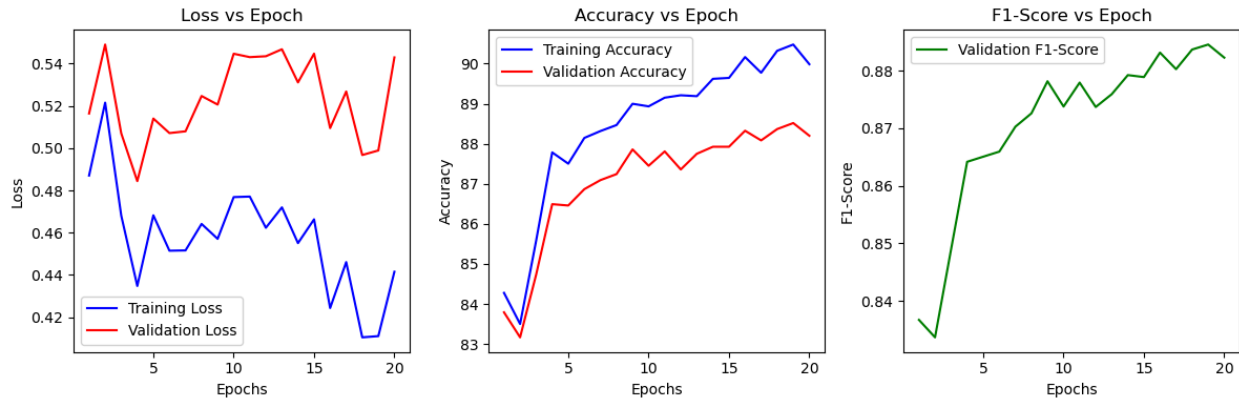
Confusion Matrix for the Best F1-Validation Score Epoch:
```
[[ 965    1   24   42    0    2   97    0   11    0]
 [   6 1174    2   15    2    0    2    0    1    0]
 [   7    0 1019   12   81    0   94    0    8    0]
 [  34    4    5 1085   31    1   19    0    3    1]
 [   1    2  107   36  946    0   77    0    4    0]
 [   3    0    0    0    0 1136    0   44    5   16]
 [ 156    2  109   25   76    0  835    0   15    0]
 [   0    0    1    0    0   26    1 1177    0   34]
 [   8    0    9    9    5    1    8    5 1175    1]
 [   0    1    0    0    0   22    0   55    0 1119]]
```

iv) Learning Rate = 0.001:

Epoch 1/20, Training Loss: 0.4871, Training Accuracy: 84.2771, Validation Loss: 0.5164, Validation Accuracy: 83.7917, Validation Macro-F1: 0.8367
Epoch 2/20, Training Loss: 0.5216, Training Accuracy: 83.5000, Validation Loss: 0.5491, Validation Accuracy: 83.1667, Validation Macro-F1: 0.8337
Epoch 3/20, Training Loss: 0.4681, Training Accuracy: 85.5812, Validation Loss: 0.5070, Validation Accuracy: 84.7333, Validation Macro-F1: 0.8490
Epoch 4/20, Training Loss: 0.4348, Training Accuracy: 87.7833, Validation Loss: 0.4844, Validation Accuracy: 86.4917, Validation Macro-F1: 0.8642
Epoch 5/20, Training Loss: 0.4682, Training Accuracy: 87.5021, Validation Loss: 0.5140, Validation Accuracy: 86.4583, Validation Macro-F1: 0.8651
Epoch 6/20, Training Loss: 0.4515, Training Accuracy: 88.1458, Validation Loss: 0.5072, Validation Accuracy: 86.8667, Validation Macro-F1: 0.8659
Epoch 7/20, Training Loss: 0.4516, Training Accuracy: 88.3167, Validation Loss: 0.5080, Validation Accuracy: 87.0917, Validation Macro-F1: 0.8703
Epoch 8/20, Training Loss: 0.4642, Training Accuracy: 88.4667, Validation Loss: 0.5247, Validation Accuracy: 87.2417, Validation Macro-F1: 0.8726
Epoch 9/20, Training Loss: 0.4571, Training Accuracy: 88.9979, Validation Loss: 0.5206, Validation Accuracy: 87.8583, Validation Macro-F1: 0.8782
Epoch 10/20, Training Loss: 0.4769, Training Accuracy: 88.9333, Validation Loss: 0.5447, Validation Accuracy: 87.4500, Validation Macro-F1: 0.8738
Epoch 11/20, Training Loss: 0.4771, Training Accuracy: 89.1500, Validation Loss: 0.5431, Validation Accuracy: 87.8083, Validation Macro-F1: 0.8779
Epoch 12/20, Training Loss: 0.4623, Training Accuracy: 89.2104, Validation Loss: 0.5435, Validation Accuracy: 87.3583, Validation Macro-F1: 0.8737
Epoch 13/20, Training Loss: 0.4720, Training Accuracy: 89.1875, Validation Loss: 0.5468, Validation Accuracy: 87.7500, Validation Macro-F1: 0.8759
Epoch 14/20, Training Loss: 0.4551, Training Accuracy: 89.6208, Validation Loss: 0.5311, Validation Accuracy: 87.9250, Validation Macro-F1: 0.8792
Epoch 15/20, Training Loss: 0.4663, Training Accuracy: 89.6437, Validation Loss: 0.5447, Validation Accuracy: 87.9250, Validation Macro-F1: 0.8789
Epoch 16/20, Training Loss: 0.4243, Training Accuracy: 90.1646, Validation Loss: 0.5095, Validation Accuracy: 88.3250, Validation Macro-F1: 0.8832
Epoch 17/20, Training Loss: 0.4461, Training Accuracy: 89.7729, Validation Loss: 0.5268, Validation Accuracy: 88.0833, Validation Macro-F1: 0.8803
Epoch 18/20, Training Loss: 0.4105, Training Accuracy: 90.3208, Validation Loss: 0.4968, Validation Accuracy: 88.3667, Validation Macro-F1: 0.8837
Epoch 19/20, Training Loss: 0.4110, Training Accuracy: 90.4771, Validation Loss: 0.4989, Validation Accuracy: 88.5167, Validation Macro-F1: 0.8846
Epoch 20/20, Training Loss: 0.4415, Training Accuracy: 89.9854, Validation Loss: 0.5430, Validation Accuracy: 88.2000, Validation Macro-F1: 0.8823

Loss vs Epoch — Accuracy vs Epoch — F1-Score vs Epoch

Confusion Matrix for the Best F1-Validation Score Epoch:

[[1033   1   15   28   9   1   99   0   18   0]
 [7 1224   3   24   3   0   5   0   1   0]
 [21   1 972   7 117   0   75   0   14   0]
 [32   9   8 1060   51   0   38   1   6   0]
 [5   2   89   25 1003   0   62   0   5   1]
 [0   0   0   1   0 1142   0   36   4   20]
 [149   0   99   19   83   0 815   1   18   0]
 [0   0   0   0   0   20   0 1119   0   34]
 [0   0   7   0   8   3   19   5 1124   5]
 [1   0   0   0   0   9   0   54   0 1130]]

Best Chosen Model:

(Xavier Initialization with 3 hidden layers, epoch = 20, batch size = 64, learning rate = 0.001)
Epoch 20/20, Training Loss: 0.3563, Training Accuracy: 91.1083, Validation Loss: 0.4353,
Validation Accuracy: 89.2667, Validation Macro-F1: 0.8922
Test Loss: 0.5548, Test Accuracy: 87.4400, Test Macro-F1: 0.8726