In [1]:

```
import tensorflow as tf
#import tensorflow_datasets as tfds
```

Baseline model

In [2]:

```
# Load Data
# tfds.list_builders()
# (train, test), info = tfds.load("mnist",split=['train', 'test'], with_info=True, as_supervised=True)
(x_train, y_train), (x_test, y_test) = tf.keras.datasets.mnist.load_data()
#(x_train, y_train), (x_test, y_test) = tf.keras.datasets.fashion_mnist.load_data()
```

```
Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datasets/mnist.npz
11493376/11490434 [==============================] - 0s 0us/step
11501568/11490434 [==============================] - 0s 0us/step
```

In [3]:

```
# Baseline model definition
model = tf.keras.Sequential([
                             tf.keras.layers.Flatten(input_shape=(28, 28)),
                             tf.keras.layers.Dense(16, activation='relu'),
                             tf.keras.layers.Dense(16, activation='relu'),
                             tf.keras.layers.Dense(10, activation='softmax')
])
```

In [4]:

```
# Baseline model compilation
model.compile(optimizer='adam',
              loss='sparse_categorical_crossentropy',
              metrics=['sparse_categorical_accuracy']
             )
```

In [5]:

```
# Baseline model fitting
history = model.fit(x_train, y_train,
                    batch_size=128,
                    epochs=100,
                    validation_data=(x_test, y_test),
                    verbose=2
                    )
```

```
Epoch 1/100
469/469 - 4s - loss: 3.4456 - sparse_categorical_accuracy: 0.1932 - val_loss: 2.0675 - val_sparse_ca
tegorical_accuracy: 0.2054 - 4s/epoch - 8ms/step
Epoch 2/100
469/469 - 2s - loss: 2.0427 - sparse_categorical_accuracy: 0.2061 - val_loss: 2.0334 - val_sparse_ca
tegorical_accuracy: 0.2078 - 2s/epoch - 4ms/step
Epoch 3/100
469/469 - 2s - loss: 2.0152 - sparse_categorical_accuracy: 0.2096 - val_loss: 2.0110 - val_sparse_ca
tegorical_accuracy: 0.2105 - 2s/epoch - 4ms/step
Epoch 4/100
469/469 - 2s - loss: 2.0011 - sparse_categorical_accuracy: 0.2094 - val_loss: 1.9966 - val_sparse_ca
tegorical_accuracy: 0.2126 - 2s/epoch - 4ms/step
Epoch 5/100
469/469 - 2s - loss: 1.9753 - sparse_categorical_accuracy: 0.2210 - val_loss: 1.9698 - val_sparse_ca
tegorical_accuracy: 0.2269 - 2s/epoch - 4ms/step
Epoch 6/100
469/469 - 1s - loss: 1.9513 - sparse_categorical_accuracy: 0.2305 - val_loss: 1.9511 - val_sparse_ca
tegorical_accuracy: 0.2342 - 1s/epoch - 3ms/step
Epoch 7/100
469/469 - 1s - loss: 1.9056 - sparse_categorical_accuracy: 0.2498 - val_loss: 1.8883 - val_sparse_ca
tegorical_accuracy: 0.2637 - 961ms/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 1.7445 - sparse_categorical_accuracy: 0.3287 - val_loss: 1.5570 - val_sparse_ca
tegorical_accuracy: 0.4081 - 945ms/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 1.4151 - sparse_categorical_accuracy: 0.4676 - val_loss: 1.2790 - val_sparse_ca
tegorical_accuracy: 0.5304 - 929ms/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 1.1866 - sparse_categorical_accuracy: 0.5476 - val_loss: 1.1090 - val_sparse_ca
tegorical_accuracy: 0.5745 - 907ms/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 1.0255 - sparse_categorical_accuracy: 0.6120 - val_loss: 0.9805 - val_sparse_ca
tegorical_accuracy: 0.6349 - 942ms/epoch - 2ms/step
```

```
Epoch 12/100
469/469 - 1s - loss: 0.9026 - sparse_categorical_accuracy: 0.6537 - val_loss: 0.8674 - val_sparse_ca
tegorical_accuracy: 0.6579 - 893ms/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.8202 - sparse_categorical_accuracy: 0.7118 - val_loss: 0.7815 - val_sparse_ca
tegorical_accuracy: 0.7533 - 952ms/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.7487 - sparse_categorical_accuracy: 0.7584 - val_loss: 0.7496 - val_sparse_ca
tegorical_accuracy: 0.7687 - 967ms/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.6008 - sparse_categorical_accuracy: 0.8244 - val_loss: 0.5650 - val_sparse_ca
tegorical_accuracy: 0.8399 - 945ms/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.5404 - sparse_categorical_accuracy: 0.8450 - val_loss: 0.5424 - val_sparse_ca
tegorical_accuracy: 0.8508 - 912ms/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.5157 - sparse_categorical_accuracy: 0.8499 - val_loss: 0.5190 - val_sparse_ca
tegorical_accuracy: 0.8502 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.4993 - sparse_categorical_accuracy: 0.8528 - val_loss: 0.5070 - val_sparse_ca
tegorical_accuracy: 0.8569 - 924ms/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.4791 - sparse_categorical_accuracy: 0.8581 - val_loss: 0.4912 - val_sparse_ca
tegorical_accuracy: 0.8545 - 915ms/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.4668 - sparse_categorical_accuracy: 0.8621 - val_loss: 0.4812 - val_sparse_ca
tegorical_accuracy: 0.8595 - 982ms/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.4516 - sparse_categorical_accuracy: 0.8670 - val_loss: 0.5094 - val_sparse_ca
tegorical_accuracy: 0.8524 - 923ms/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.4410 - sparse_categorical_accuracy: 0.8710 - val_loss: 0.4602 - val_sparse_ca
tegorical_accuracy: 0.8704 - 911ms/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.4346 - sparse_categorical_accuracy: 0.8749 - val_loss: 0.4755 - val_sparse_ca
tegorical_accuracy: 0.8687 - 912ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.4209 - sparse_categorical_accuracy: 0.8774 - val_loss: 0.4358 - val_sparse_ca
tegorical_accuracy: 0.8790 - 980ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.4155 - sparse_categorical_accuracy: 0.8796 - val_loss: 0.4383 - val_sparse_ca
tegorical_accuracy: 0.8751 - 897ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.4079 - sparse_categorical_accuracy: 0.8808 - val_loss: 0.4496 - val_sparse_ca
tegorical_accuracy: 0.8756 - 948ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.4049 - sparse_categorical_accuracy: 0.8815 - val_loss: 0.4453 - val_sparse_ca
tegorical_accuracy: 0.8780 - 976ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.4012 - sparse_categorical_accuracy: 0.8820 - val_loss: 0.4422 - val_sparse_ca
tegorical_accuracy: 0.8772 - 931ms/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.4027 - sparse_categorical_accuracy: 0.8819 - val_loss: 0.4430 - val_sparse_ca
tegorical_accuracy: 0.8766 - 933ms/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.3983 - sparse_categorical_accuracy: 0.8821 - val_loss: 0.4251 - val_sparse_ca
tegorical_accuracy: 0.8802 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.3925 - sparse_categorical_accuracy: 0.8860 - val_loss: 0.4491 - val_sparse_ca
tegorical_accuracy: 0.8725 - 906ms/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.3920 - sparse_categorical_accuracy: 0.8857 - val_loss: 0.4222 - val_sparse_ca
tegorical_accuracy: 0.8826 - 954ms/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.3893 - sparse_categorical_accuracy: 0.8865 - val_loss: 0.4262 - val_sparse_ca
tegorical_accuracy: 0.8821 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.3885 - sparse_categorical_accuracy: 0.8867 - val_loss: 0.4428 - val_sparse_ca
tegorical_accuracy: 0.8755 - 967ms/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.3839 - sparse_categorical_accuracy: 0.8868 - val_loss: 0.4421 - val_sparse_ca
tegorical_accuracy: 0.8770 - 966ms/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.3832 - sparse_categorical_accuracy: 0.8873 - val_loss: 0.4250 - val_sparse_ca
tegorical_accuracy: 0.8770 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.3785 - sparse_categorical_accuracy: 0.8894 - val_loss: 0.4270 - val_sparse_ca
tegorical_accuracy: 0.8790 - 948ms/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.3787 - sparse_categorical_accuracy: 0.8888 - val_loss: 0.4419 - val_sparse_ca
tegorical_accuracy: 0.8768 - 916ms/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.3778 - sparse_categorical_accuracy: 0.8886 - val_loss: 0.4331 - val_sparse_ca
```

```
tegorical_accuracy: 0.8791 - 908ms/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.3753 - sparse_categorical_accuracy: 0.8894 - val_loss: 0.4236 - val_sparse_ca
tegorical_accuracy: 0.8792 - 959ms/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.3711 - sparse_categorical_accuracy: 0.8902 - val_loss: 0.4492 - val_sparse_ca
tegorical_accuracy: 0.8696 - 947ms/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.3748 - sparse_categorical_accuracy: 0.8905 - val_loss: 0.4278 - val_sparse_ca
tegorical_accuracy: 0.8807 - 966ms/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.3717 - sparse_categorical_accuracy: 0.8914 - val_loss: 0.4207 - val_sparse_ca
tegorical_accuracy: 0.8830 - 964ms/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.3707 - sparse_categorical_accuracy: 0.8907 - val_loss: 0.4355 - val_sparse_ca
tegorical_accuracy: 0.8793 - 982ms/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.3662 - sparse_categorical_accuracy: 0.8928 - val_loss: 0.4296 - val_sparse_ca
tegorical_accuracy: 0.8771 - 945ms/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.3669 - sparse_categorical_accuracy: 0.8920 - val_loss: 0.4294 - val_sparse_ca
tegorical_accuracy: 0.8789 - 994ms/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.3666 - sparse_categorical_accuracy: 0.8911 - val_loss: 0.4231 - val_sparse_ca
tegorical_accuracy: 0.8836 - 971ms/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.3611 - sparse_categorical_accuracy: 0.8937 - val_loss: 0.4325 - val_sparse_ca
tegorical_accuracy: 0.8741 - 981ms/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.3565 - sparse_categorical_accuracy: 0.8954 - val_loss: 0.4148 - val_sparse_ca
tegorical_accuracy: 0.8866 - 905ms/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.3567 - sparse_categorical_accuracy: 0.8949 - val_loss: 0.4275 - val_sparse_ca
tegorical_accuracy: 0.8801 - 964ms/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.3575 - sparse_categorical_accuracy: 0.8953 - val_loss: 0.4151 - val_sparse_ca
tegorical_accuracy: 0.8864 - 941ms/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.3520 - sparse_categorical_accuracy: 0.8964 - val_loss: 0.4027 - val_sparse_ca
tegorical_accuracy: 0.8854 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.3511 - sparse_categorical_accuracy: 0.8961 - val_loss: 0.4273 - val_sparse_ca
tegorical_accuracy: 0.8841 - 963ms/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.3512 - sparse_categorical_accuracy: 0.8964 - val_loss: 0.4081 - val_sparse_ca
tegorical_accuracy: 0.8842 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.3498 - sparse_categorical_accuracy: 0.8962 - val_loss: 0.4050 - val_sparse_ca
tegorical_accuracy: 0.8874 - 948ms/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.3469 - sparse_categorical_accuracy: 0.8984 - val_loss: 0.4187 - val_sparse_ca
tegorical_accuracy: 0.8874 - 968ms/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.3465 - sparse_categorical_accuracy: 0.8981 - val_loss: 0.4075 - val_sparse_ca
tegorical_accuracy: 0.8838 - 910ms/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.3458 - sparse_categorical_accuracy: 0.8972 - val_loss: 0.4048 - val_sparse_ca
tegorical_accuracy: 0.8881 - 958ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.3386 - sparse_categorical_accuracy: 0.9001 - val_loss: 0.3930 - val_sparse_ca
tegorical_accuracy: 0.8902 - 957ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.3370 - sparse_categorical_accuracy: 0.9006 - val_loss: 0.3987 - val_sparse_ca
tegorical_accuracy: 0.8885 - 958ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.3374 - sparse_categorical_accuracy: 0.9016 - val_loss: 0.3768 - val_sparse_ca
tegorical_accuracy: 0.8923 - 919ms/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.3325 - sparse_categorical_accuracy: 0.9036 - val_loss: 0.3728 - val_sparse_ca
tegorical_accuracy: 0.8974 - 956ms/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.3297 - sparse_categorical_accuracy: 0.9050 - val_loss: 0.3730 - val_sparse_ca
tegorical_accuracy: 0.8988 - 944ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.3243 - sparse_categorical_accuracy: 0.9073 - val_loss: 0.3828 - val_sparse_ca
tegorical_accuracy: 0.8968 - 997ms/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.3262 - sparse_categorical_accuracy: 0.9068 - val_loss: 0.3733 - val_sparse_ca
tegorical_accuracy: 0.8990 - 918ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.3203 - sparse_categorical_accuracy: 0.9089 - val_loss: 0.3668 - val_sparse_ca
tegorical_accuracy: 0.9012 - 908ms/epoch - 2ms/step
Epoch 67/100
```

```
469/469 - 1s - loss: 0.3157 - sparse_categorical_accuracy: 0.9098 - val_loss: 0.3612 - val_sparse_ca
tegorical_accuracy: 0.8998 - 957ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.3102 - sparse_categorical_accuracy: 0.9126 - val_loss: 0.3753 - val_sparse_ca
tegorical_accuracy: 0.8986 - 924ms/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.3058 - sparse_categorical_accuracy: 0.9133 - val_loss: 0.3451 - val_sparse_ca
tegorical_accuracy: 0.9064 - 937ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.3065 - sparse_categorical_accuracy: 0.9140 - val_loss: 0.3536 - val_sparse_ca
tegorical_accuracy: 0.9036 - 964ms/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.3016 - sparse_categorical_accuracy: 0.9154 - val_loss: 0.3562 - val_sparse_ca
tegorical_accuracy: 0.9063 - 913ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.3024 - sparse_categorical_accuracy: 0.9148 - val_loss: 0.3446 - val_sparse_ca
tegorical_accuracy: 0.9094 - 915ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.2952 - sparse_categorical_accuracy: 0.9174 - val_loss: 0.3640 - val_sparse_ca
tegorical_accuracy: 0.9069 - 947ms/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.2949 - sparse_categorical_accuracy: 0.9180 - val_loss: 0.3362 - val_sparse_ca
tegorical_accuracy: 0.9103 - 915ms/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.2959 - sparse_categorical_accuracy: 0.9165 - val_loss: 0.3396 - val_sparse_ca
tegorical_accuracy: 0.9111 - 961ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.2921 - sparse_categorical_accuracy: 0.9180 - val_loss: 0.3464 - val_sparse_ca
tegorical_accuracy: 0.9108 - 934ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.2910 - sparse_categorical_accuracy: 0.9183 - val_loss: 0.3488 - val_sparse_ca
tegorical_accuracy: 0.9074 - 963ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.2879 - sparse_categorical_accuracy: 0.9197 - val_loss: 0.3381 - val_sparse_ca
tegorical_accuracy: 0.9133 - 962ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.2866 - sparse_categorical_accuracy: 0.9196 - val_loss: 0.3448 - val_sparse_ca
tegorical_accuracy: 0.9115 - 964ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.2863 - sparse_categorical_accuracy: 0.9179 - val_loss: 0.3301 - val_sparse_ca
tegorical_accuracy: 0.9127 - 950ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.2853 - sparse_categorical_accuracy: 0.9196 - val_loss: 0.3368 - val_sparse_ca
tegorical_accuracy: 0.9135 - 960ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.2849 - sparse_categorical_accuracy: 0.9203 - val_loss: 0.3482 - val_sparse_ca
tegorical_accuracy: 0.9080 - 967ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.2820 - sparse_categorical_accuracy: 0.9206 - val_loss: 0.3386 - val_sparse_ca
tegorical_accuracy: 0.9114 - 1s/epoch - 3ms/step
Epoch 84/100
469/469 - 1s - loss: 0.2830 - sparse_categorical_accuracy: 0.9205 - val_loss: 0.3407 - val_sparse_ca
tegorical_accuracy: 0.9101 - 1s/epoch - 3ms/step
Epoch 85/100
469/469 - 1s - loss: 0.2815 - sparse_categorical_accuracy: 0.9213 - val_loss: 0.3372 - val_sparse_ca
tegorical_accuracy: 0.9114 - 1s/epoch - 3ms/step
Epoch 86/100
469/469 - 1s - loss: 0.2799 - sparse_categorical_accuracy: 0.9213 - val_loss: 0.3347 - val_sparse_ca
tegorical_accuracy: 0.9108 - 909ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.2772 - sparse_categorical_accuracy: 0.9227 - val_loss: 0.3513 - val_sparse_ca
tegorical_accuracy: 0.9081 - 993ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.2796 - sparse_categorical_accuracy: 0.9206 - val_loss: 0.3430 - val_sparse_ca
tegorical_accuracy: 0.9126 - 959ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.2771 - sparse_categorical_accuracy: 0.9214 - val_loss: 0.3357 - val_sparse_ca
tegorical_accuracy: 0.9148 - 918ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.2778 - sparse_categorical_accuracy: 0.9216 - val_loss: 0.3334 - val_sparse_ca
tegorical_accuracy: 0.9153 - 975ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.2750 - sparse_categorical_accuracy: 0.9217 - val_loss: 0.3390 - val_sparse_ca
tegorical_accuracy: 0.9092 - 912ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.2755 - sparse_categorical_accuracy: 0.9220 - val_loss: 0.3334 - val_sparse_ca
tegorical_accuracy: 0.9149 - 911ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.2728 - sparse_categorical_accuracy: 0.9234 - val_loss: 0.3379 - val_sparse_ca
tegorical_accuracy: 0.9114 - 907ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.2720 - sparse_categorical_accuracy: 0.9224 - val_loss: 0.3435 - val_sparse_ca
tegorical_accuracy: 0.9128 - 972ms/epoch - 2ms/step
```

```
Epoch 95/100
469/469 - 1s - loss: 0.2723 - sparse_categorical_accuracy: 0.9232 - val_loss: 0.3455 - val_sparse_ca
tegorical_accuracy: 0.9102 - 918ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.2726 - sparse_categorical_accuracy: 0.9227 - val_loss: 0.3328 - val_sparse_ca
tegorical_accuracy: 0.9129 - 952ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.2732 - sparse_categorical_accuracy: 0.9226 - val_loss: 0.3618 - val_sparse_ca
tegorical_accuracy: 0.9029 - 920ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.2701 - sparse_categorical_accuracy: 0.9239 - val_loss: 0.3331 - val_sparse_ca
tegorical_accuracy: 0.9126 - 971ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.2706 - sparse_categorical_accuracy: 0.9233 - val_loss: 0.3481 - val_sparse_ca
tegorical_accuracy: 0.9084 - 928ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.2690 - sparse_categorical_accuracy: 0.9238 - val_loss: 0.3364 - val_sparse_ca
tegorical_accuracy: 0.9144 - 914ms/epoch - 2ms/step
```

In [6]:

```
# Baseline model evaluation
model.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.3364 - sparse_categorical_accuracy: 0.9144 - 348ms/epoch - 1ms/step
```

Out[6]:

```
[0.33644601702690125, 0.9143999814987183]
```

In [7]:

```
import keras
from tensorflow.keras.layers import Layer
import keras.backend as K
import tensorflow as tf
from keras.datasets import cifar10
from keras.models import Model
from keras.layers import Conv2D, MaxPool2D,  \
    Dropout, Dense, Input, concatenate,      \
    GlobalAveragePooling2D, AveragePooling2D,\
    Flatten

import cv2
import numpy as np
from keras.datasets import cifar10
from keras import backend as K
#from keras.utils import np_utils
kernel_init = keras.initializers.glorot_uniform()
bias_init = keras.initializers.Constant(value=0.2)
```

In [8]:

```
# Inception V3 Module
def inception_module(x,
                     filters_1x1,
                     filters_3x3_reduce,
                     filters_3x3,
                     filters_5x5_reduce,
                     filters_5x5,
                     filters_pool_proj,
                     name=None):

    conv_1x1 = Conv2D(filters_1x1, (1, 1), padding='same', activation='relu', kernel_initializer=kernel_init, bia
s_initializer=bias_init)(x)

    conv_3x3 = Conv2D(filters_3x3_reduce, (1, 1), padding='same', activation='relu', kernel_initializer=kernel_in
it, bias_initializer=bias_init)(x)
    conv_3x3 = Conv2D(filters_3x3, (3, 3), padding='same', activation='relu', kernel_initializer=kernel_init, bia
s_initializer=bias_init)(conv_3x3)

    conv_5x5 = Conv2D(filters_5x5_reduce, (1, 1), padding='same', activation='relu', kernel_initializer=kernel_in
it, bias_initializer=bias_init)(x)
    conv_5x5 = Conv2D(filters_5x5, (5, 5), padding='same', activation='relu', kernel_initializer=kernel_init, bia
s_initializer=bias_init)(conv_5x5)

    pool_proj = MaxPool2D((3, 3), strides=(1, 1), padding='same')(x)
    pool_proj = Conv2D(filters_pool_proj, (1, 1), padding='same', activation='relu', kernel_initializer=kernel_in
it, bias_initializer=bias_init)(pool_proj)

    output = concatenate([conv_1x1, conv_3x3, conv_5x5, pool_proj], axis=1, name=name)

    return output
```

The baseline model with `kernel_initializer='glorot_uniform'`, `bias_initializer='zeros'`

Try different models

Model1: `adam` optimizer with learning rate=e^{-3}, `random_uniform` initializer, `dropout` regularization with rate=0.1.

In [9]:

```
input_layer = Input(shape=(28, 28,1))
x = inception_module(input_layer,
                     filters_1x1=1,
                     filters_3x3_reduce=1,
                     filters_3x3=1,
                     filters_5x5_reduce=1,
                     filters_5x5=1,
                     filters_pool_proj=1,
                     name='inception_3a')
x = tf.keras.layers.BatchNormalization()(x)
x = MaxPool2D((3, 3))(x)
x = tf.keras.layers.Flatten()(x)
x = Dropout(0.1)(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform')(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform')(x)
x = tf.keras.layers.BatchNormalization()(x)
x = Dense(10, activation='softmax', kernel_initializer='random_uniform')(x)
model1 = Model(input_layer, x, name='inception_v1')
```

In [10]:

```
model1.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [11]:

```
history1 = model1.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 53s - loss: 0.9712 - sparse_categorical_accuracy: 0.7746 - val_loss: 0.3838 - val_sparse_c
ategorical_accuracy: 0.8945 - 53s/epoch - 113ms/step
Epoch 2/100
469/469 - 52s - loss: 0.4190 - sparse_categorical_accuracy: 0.8711 - val_loss: 0.3234 - val_sparse_c
ategorical_accuracy: 0.8981 - 52s/epoch - 111ms/step
Epoch 3/100
469/469 - 52s - loss: 0.3716 - sparse_categorical_accuracy: 0.8818 - val_loss: 0.2864 - val_sparse_c
ategorical_accuracy: 0.9110 - 52s/epoch - 111ms/step
Epoch 4/100
469/469 - 52s - loss: 0.3517 - sparse_categorical_accuracy: 0.8872 - val_loss: 0.2683 - val_sparse_c
ategorical_accuracy: 0.9155 - 52s/epoch - 111ms/step
Epoch 5/100
469/469 - 52s - loss: 0.3205 - sparse_categorical_accuracy: 0.8969 - val_loss: 0.2441 - val_sparse_c
ategorical_accuracy: 0.9243 - 52s/epoch - 111ms/step
Epoch 6/100
469/469 - 53s - loss: 0.2871 - sparse_categorical_accuracy: 0.9079 - val_loss: 0.2521 - val_sparse_c
ategorical_accuracy: 0.9205 - 53s/epoch - 112ms/step
Epoch 7/100
469/469 - 52s - loss: 0.2627 - sparse_categorical_accuracy: 0.9172 - val_loss: 0.2014 - val_sparse_c
ategorical_accuracy: 0.9372 - 52s/epoch - 111ms/step
Epoch 8/100
469/469 - 52s - loss: 0.2454 - sparse_categorical_accuracy: 0.9205 - val_loss: 0.1922 - val_sparse_c
ategorical_accuracy: 0.9413 - 52s/epoch - 111ms/step
Epoch 9/100
469/469 - 52s - loss: 0.2313 - sparse_categorical_accuracy: 0.9273 - val_loss: 0.1949 - val_sparse_c
ategorical_accuracy: 0.9399 - 52s/epoch - 112ms/step
Epoch 10/100
469/469 - 52s - loss: 0.2192 - sparse_categorical_accuracy: 0.9306 - val_loss: 0.1693 - val_sparse_c
ategorical_accuracy: 0.9475 - 52s/epoch - 111ms/step
Epoch 11/100
469/469 - 52s - loss: 0.2111 - sparse_categorical_accuracy: 0.9334 - val_loss: 0.1765 - val_sparse_c
ategorical_accuracy: 0.9451 - 52s/epoch - 111ms/step
Epoch 12/100
469/469 - 52s - loss: 0.2023 - sparse_categorical_accuracy: 0.9356 - val_loss: 0.1601 - val_sparse_c
ategorical_accuracy: 0.9513 - 52s/epoch - 111ms/step
```

```
Epoch 13/100
469/469 - 52s - loss: 0.1945 - sparse_categorical_accuracy: 0.9388 - val_loss: 0.1553 - val_sparse_c
ategorical_accuracy: 0.9521 - 52s/epoch - 111ms/step
Epoch 14/100
469/469 - 52s - loss: 0.1915 - sparse_categorical_accuracy: 0.9390 - val_loss: 0.1602 - val_sparse_c
ategorical_accuracy: 0.9507 - 52s/epoch - 112ms/step
Epoch 15/100
469/469 - 52s - loss: 0.1866 - sparse_categorical_accuracy: 0.9406 - val_loss: 0.1400 - val_sparse_c
ategorical_accuracy: 0.9563 - 52s/epoch - 111ms/step
Epoch 16/100
469/469 - 52s - loss: 0.1826 - sparse_categorical_accuracy: 0.9426 - val_loss: 0.1455 - val_sparse_c
ategorical_accuracy: 0.9539 - 52s/epoch - 112ms/step
Epoch 17/100
469/469 - 52s - loss: 0.1779 - sparse_categorical_accuracy: 0.9431 - val_loss: 0.1481 - val_sparse_c
ategorical_accuracy: 0.9543 - 52s/epoch - 111ms/step
Epoch 18/100
469/469 - 52s - loss: 0.1746 - sparse_categorical_accuracy: 0.9443 - val_loss: 0.1395 - val_sparse_c
ategorical_accuracy: 0.9570 - 52s/epoch - 112ms/step
Epoch 19/100
469/469 - 52s - loss: 0.1723 - sparse_categorical_accuracy: 0.9451 - val_loss: 0.1395 - val_sparse_c
ategorical_accuracy: 0.9564 - 52s/epoch - 112ms/step
Epoch 20/100
469/469 - 52s - loss: 0.1719 - sparse_categorical_accuracy: 0.9464 - val_loss: 0.1361 - val_sparse_c
ategorical_accuracy: 0.9564 - 52s/epoch - 112ms/step
Epoch 21/100
469/469 - 53s - loss: 0.1694 - sparse_categorical_accuracy: 0.9470 - val_loss: 0.1392 - val_sparse_c
ategorical_accuracy: 0.9550 - 53s/epoch - 112ms/step
Epoch 22/100
469/469 - 53s - loss: 0.1662 - sparse_categorical_accuracy: 0.9473 - val_loss: 0.1405 - val_sparse_c
ategorical_accuracy: 0.9564 - 53s/epoch - 113ms/step
Epoch 23/100
469/469 - 52s - loss: 0.1666 - sparse_categorical_accuracy: 0.9463 - val_loss: 0.1463 - val_sparse_c
ategorical_accuracy: 0.9531 - 52s/epoch - 112ms/step
Epoch 24/100
469/469 - 53s - loss: 0.1628 - sparse_categorical_accuracy: 0.9487 - val_loss: 0.1349 - val_sparse_c
ategorical_accuracy: 0.9572 - 53s/epoch - 112ms/step
Epoch 25/100
469/469 - 53s - loss: 0.1632 - sparse_categorical_accuracy: 0.9488 - val_loss: 0.1367 - val_sparse_c
ategorical_accuracy: 0.9563 - 53s/epoch - 112ms/step
Epoch 26/100
469/469 - 53s - loss: 0.1615 - sparse_categorical_accuracy: 0.9499 - val_loss: 0.1238 - val_sparse_c
ategorical_accuracy: 0.9621 - 53s/epoch - 112ms/step
Epoch 27/100
469/469 - 53s - loss: 0.1600 - sparse_categorical_accuracy: 0.9498 - val_loss: 0.1305 - val_sparse_c
ategorical_accuracy: 0.9590 - 53s/epoch - 112ms/step
Epoch 28/100
469/469 - 53s - loss: 0.1564 - sparse_categorical_accuracy: 0.9498 - val_loss: 0.1209 - val_sparse_c
ategorical_accuracy: 0.9599 - 53s/epoch - 113ms/step
Epoch 29/100
469/469 - 53s - loss: 0.1541 - sparse_categorical_accuracy: 0.9515 - val_loss: 0.1218 - val_sparse_c
ategorical_accuracy: 0.9605 - 53s/epoch - 113ms/step
Epoch 30/100
469/469 - 53s - loss: 0.1547 - sparse_categorical_accuracy: 0.9509 - val_loss: 0.1379 - val_sparse_c
ategorical_accuracy: 0.9553 - 53s/epoch - 113ms/step
Epoch 31/100
469/469 - 53s - loss: 0.1524 - sparse_categorical_accuracy: 0.9526 - val_loss: 0.1259 - val_sparse_c
ategorical_accuracy: 0.9600 - 53s/epoch - 113ms/step
Epoch 32/100
469/469 - 53s - loss: 0.1500 - sparse_categorical_accuracy: 0.9525 - val_loss: 0.1164 - val_sparse_c
ategorical_accuracy: 0.9629 - 53s/epoch - 113ms/step
Epoch 33/100
469/469 - 53s - loss: 0.1491 - sparse_categorical_accuracy: 0.9530 - val_loss: 0.1124 - val_sparse_c
ategorical_accuracy: 0.9629 - 53s/epoch - 112ms/step
Epoch 34/100
469/469 - 53s - loss: 0.1481 - sparse_categorical_accuracy: 0.9534 - val_loss: 0.1442 - val_sparse_c
ategorical_accuracy: 0.9537 - 53s/epoch - 112ms/step
Epoch 35/100
469/469 - 53s - loss: 0.1468 - sparse_categorical_accuracy: 0.9538 - val_loss: 0.1218 - val_sparse_c
ategorical_accuracy: 0.9622 - 53s/epoch - 112ms/step
Epoch 36/100
469/469 - 53s - loss: 0.1439 - sparse_categorical_accuracy: 0.9546 - val_loss: 0.1135 - val_sparse_c
ategorical_accuracy: 0.9630 - 53s/epoch - 112ms/step
Epoch 37/100
469/469 - 52s - loss: 0.1438 - sparse_categorical_accuracy: 0.9552 - val_loss: 0.1159 - val_sparse_c
ategorical_accuracy: 0.9622 - 52s/epoch - 112ms/step
Epoch 38/100
469/469 - 52s - loss: 0.1398 - sparse_categorical_accuracy: 0.9557 - val_loss: 0.1105 - val_sparse_c
ategorical_accuracy: 0.9642 - 52s/epoch - 111ms/step
Epoch 39/100
469/469 - 52s - loss: 0.1407 - sparse_categorical_accuracy: 0.9559 - val_loss: 0.1183 - val_sparse_c
ategorical_accuracy: 0.9604 - 52s/epoch - 111ms/step
Epoch 40/100
469/469 - 52s - loss: 0.1395 - sparse_categorical_accuracy: 0.9561 - val_loss: 0.1101 - val_sparse_c
```

```
ategorical_accuracy: 0.9632 - 52s/epoch - 112ms/step
Epoch 41/100
469/469 - 52s - loss: 0.1387 - sparse_categorical_accuracy: 0.9557 - val_loss: 0.1086 - val_sparse_c
ategorical_accuracy: 0.9652 - 52s/epoch - 111ms/step
Epoch 42/100
469/469 - 52s - loss: 0.1360 - sparse_categorical_accuracy: 0.9561 - val_loss: 0.1127 - val_sparse_c
ategorical_accuracy: 0.9636 - 52s/epoch - 112ms/step
Epoch 43/100
469/469 - 52s - loss: 0.1353 - sparse_categorical_accuracy: 0.9573 - val_loss: 0.1160 - val_sparse_c
ategorical_accuracy: 0.9642 - 52s/epoch - 112ms/step
Epoch 44/100
469/469 - 52s - loss: 0.1346 - sparse_categorical_accuracy: 0.9577 - val_loss: 0.1196 - val_sparse_c
ategorical_accuracy: 0.9626 - 52s/epoch - 112ms/step
Epoch 45/100
469/469 - 52s - loss: 0.1336 - sparse_categorical_accuracy: 0.9575 - val_loss: 0.1001 - val_sparse_c
ategorical_accuracy: 0.9681 - 52s/epoch - 111ms/step
Epoch 46/100
469/469 - 52s - loss: 0.1299 - sparse_categorical_accuracy: 0.9586 - val_loss: 0.1021 - val_sparse_c
ategorical_accuracy: 0.9665 - 52s/epoch - 112ms/step
Epoch 47/100
469/469 - 52s - loss: 0.1307 - sparse_categorical_accuracy: 0.9593 - val_loss: 0.1149 - val_sparse_c
ategorical_accuracy: 0.9640 - 52s/epoch - 112ms/step
Epoch 48/100
469/469 - 52s - loss: 0.1293 - sparse_categorical_accuracy: 0.9592 - val_loss: 0.1043 - val_sparse_c
ategorical_accuracy: 0.9671 - 52s/epoch - 111ms/step
Epoch 49/100
469/469 - 52s - loss: 0.1308 - sparse_categorical_accuracy: 0.9583 - val_loss: 0.1057 - val_sparse_c
ategorical_accuracy: 0.9648 - 52s/epoch - 111ms/step
Epoch 50/100
469/469 - 52s - loss: 0.1291 - sparse_categorical_accuracy: 0.9595 - val_loss: 0.1050 - val_sparse_c
ategorical_accuracy: 0.9667 - 52s/epoch - 111ms/step
Epoch 51/100
469/469 - 52s - loss: 0.1272 - sparse_categorical_accuracy: 0.9607 - val_loss: 0.1109 - val_sparse_c
ategorical_accuracy: 0.9649 - 52s/epoch - 111ms/step
Epoch 52/100
469/469 - 52s - loss: 0.1289 - sparse_categorical_accuracy: 0.9596 - val_loss: 0.1020 - val_sparse_c
ategorical_accuracy: 0.9682 - 52s/epoch - 111ms/step
Epoch 53/100
469/469 - 52s - loss: 0.1289 - sparse_categorical_accuracy: 0.9598 - val_loss: 0.1083 - val_sparse_c
ategorical_accuracy: 0.9649 - 52s/epoch - 111ms/step
Epoch 54/100
469/469 - 52s - loss: 0.1290 - sparse_categorical_accuracy: 0.9592 - val_loss: 0.1026 - val_sparse_c
ategorical_accuracy: 0.9686 - 52s/epoch - 111ms/step
Epoch 55/100
469/469 - 52s - loss: 0.1274 - sparse_categorical_accuracy: 0.9597 - val_loss: 0.1065 - val_sparse_c
ategorical_accuracy: 0.9662 - 52s/epoch - 111ms/step
Epoch 56/100
469/469 - 52s - loss: 0.1278 - sparse_categorical_accuracy: 0.9591 - val_loss: 0.1008 - val_sparse_c
ategorical_accuracy: 0.9695 - 52s/epoch - 111ms/step
Epoch 57/100
469/469 - 52s - loss: 0.1259 - sparse_categorical_accuracy: 0.9604 - val_loss: 0.1035 - val_sparse_c
ategorical_accuracy: 0.9663 - 52s/epoch - 111ms/step
Epoch 58/100
469/469 - 52s - loss: 0.1249 - sparse_categorical_accuracy: 0.9602 - val_loss: 0.1070 - val_sparse_c
ategorical_accuracy: 0.9652 - 52s/epoch - 112ms/step
Epoch 59/100
469/469 - 52s - loss: 0.1244 - sparse_categorical_accuracy: 0.9596 - val_loss: 0.1032 - val_sparse_c
ategorical_accuracy: 0.9670 - 52s/epoch - 111ms/step
Epoch 60/100
469/469 - 52s - loss: 0.1235 - sparse_categorical_accuracy: 0.9611 - val_loss: 0.1003 - val_sparse_c
ategorical_accuracy: 0.9684 - 52s/epoch - 111ms/step
Epoch 61/100
469/469 - 52s - loss: 0.1207 - sparse_categorical_accuracy: 0.9617 - val_loss: 0.1016 - val_sparse_c
ategorical_accuracy: 0.9674 - 52s/epoch - 111ms/step
Epoch 62/100
469/469 - 52s - loss: 0.1187 - sparse_categorical_accuracy: 0.9625 - val_loss: 0.1004 - val_sparse_c
ategorical_accuracy: 0.9677 - 52s/epoch - 111ms/step
Epoch 63/100
469/469 - 52s - loss: 0.1216 - sparse_categorical_accuracy: 0.9613 - val_loss: 0.1086 - val_sparse_c
ategorical_accuracy: 0.9655 - 52s/epoch - 111ms/step
Epoch 64/100
469/469 - 52s - loss: 0.1215 - sparse_categorical_accuracy: 0.9611 - val_loss: 0.1069 - val_sparse_c
ategorical_accuracy: 0.9660 - 52s/epoch - 111ms/step
Epoch 65/100
469/469 - 52s - loss: 0.1202 - sparse_categorical_accuracy: 0.9619 - val_loss: 0.1060 - val_sparse_c
ategorical_accuracy: 0.9652 - 52s/epoch - 111ms/step
Epoch 66/100
469/469 - 52s - loss: 0.1203 - sparse_categorical_accuracy: 0.9617 - val_loss: 0.1071 - val_sparse_c
ategorical_accuracy: 0.9650 - 52s/epoch - 111ms/step
Epoch 67/100
469/469 - 52s - loss: 0.1194 - sparse_categorical_accuracy: 0.9623 - val_loss: 0.1016 - val_sparse_c
ategorical_accuracy: 0.9685 - 52s/epoch - 111ms/step
Epoch 68/100
```

```
469/469 - 52s - loss: 0.1174 - sparse_categorical_accuracy: 0.9625 - val_loss: 0.0997 - val_sparse_c
ategorical_accuracy: 0.9687 - 52s/epoch - 111ms/step
Epoch 69/100
469/469 - 52s - loss: 0.1181 - sparse_categorical_accuracy: 0.9626 - val_loss: 0.0922 - val_sparse_c
ategorical_accuracy: 0.9703 - 52s/epoch - 111ms/step
Epoch 70/100
469/469 - 52s - loss: 0.1197 - sparse_categorical_accuracy: 0.9624 - val_loss: 0.1051 - val_sparse_c
ategorical_accuracy: 0.9656 - 52s/epoch - 111ms/step
Epoch 71/100
469/469 - 52s - loss: 0.1208 - sparse_categorical_accuracy: 0.9618 - val_loss: 0.1007 - val_sparse_c
ategorical_accuracy: 0.9678 - 52s/epoch - 111ms/step
Epoch 72/100
469/469 - 52s - loss: 0.1188 - sparse_categorical_accuracy: 0.9630 - val_loss: 0.0944 - val_sparse_c
ategorical_accuracy: 0.9700 - 52s/epoch - 111ms/step
Epoch 73/100
469/469 - 52s - loss: 0.1167 - sparse_categorical_accuracy: 0.9632 - val_loss: 0.0971 - val_sparse_c
ategorical_accuracy: 0.9689 - 52s/epoch - 111ms/step
Epoch 74/100
469/469 - 52s - loss: 0.1165 - sparse_categorical_accuracy: 0.9635 - val_loss: 0.1063 - val_sparse_c
ategorical_accuracy: 0.9650 - 52s/epoch - 111ms/step
Epoch 75/100
469/469 - 52s - loss: 0.1156 - sparse_categorical_accuracy: 0.9634 - val_loss: 0.1040 - val_sparse_c
ategorical_accuracy: 0.9677 - 52s/epoch - 112ms/step
Epoch 76/100
469/469 - 52s - loss: 0.1147 - sparse_categorical_accuracy: 0.9631 - val_loss: 0.1007 - val_sparse_c
ategorical_accuracy: 0.9691 - 52s/epoch - 111ms/step
Epoch 77/100
469/469 - 52s - loss: 0.1175 - sparse_categorical_accuracy: 0.9630 - val_loss: 0.0971 - val_sparse_c
ategorical_accuracy: 0.9702 - 52s/epoch - 111ms/step
Epoch 78/100
469/469 - 52s - loss: 0.1167 - sparse_categorical_accuracy: 0.9620 - val_loss: 0.0967 - val_sparse_c
ategorical_accuracy: 0.9700 - 52s/epoch - 111ms/step
Epoch 79/100
469/469 - 52s - loss: 0.1161 - sparse_categorical_accuracy: 0.9630 - val_loss: 0.0954 - val_sparse_c
ategorical_accuracy: 0.9703 - 52s/epoch - 111ms/step
Epoch 80/100
469/469 - 52s - loss: 0.1141 - sparse_categorical_accuracy: 0.9637 - val_loss: 0.1002 - val_sparse_c
ategorical_accuracy: 0.9689 - 52s/epoch - 111ms/step
Epoch 81/100
469/469 - 52s - loss: 0.1142 - sparse_categorical_accuracy: 0.9632 - val_loss: 0.0994 - val_sparse_c
ategorical_accuracy: 0.9686 - 52s/epoch - 112ms/step
Epoch 82/100
469/469 - 52s - loss: 0.1107 - sparse_categorical_accuracy: 0.9645 - val_loss: 0.0953 - val_sparse_c
ategorical_accuracy: 0.9703 - 52s/epoch - 111ms/step
Epoch 83/100
469/469 - 52s - loss: 0.1136 - sparse_categorical_accuracy: 0.9633 - val_loss: 0.0928 - val_sparse_c
ategorical_accuracy: 0.9720 - 52s/epoch - 111ms/step
Epoch 84/100
469/469 - 52s - loss: 0.1118 - sparse_categorical_accuracy: 0.9651 - val_loss: 0.0978 - val_sparse_c
ategorical_accuracy: 0.9695 - 52s/epoch - 111ms/step
Epoch 85/100
469/469 - 52s - loss: 0.1120 - sparse_categorical_accuracy: 0.9646 - val_loss: 0.0979 - val_sparse_c
ategorical_accuracy: 0.9703 - 52s/epoch - 111ms/step
Epoch 86/100
469/469 - 52s - loss: 0.1115 - sparse_categorical_accuracy: 0.9642 - val_loss: 0.0916 - val_sparse_c
ategorical_accuracy: 0.9715 - 52s/epoch - 111ms/step
Epoch 87/100
469/469 - 52s - loss: 0.1128 - sparse_categorical_accuracy: 0.9638 - val_loss: 0.0955 - val_sparse_c
ategorical_accuracy: 0.9713 - 52s/epoch - 111ms/step
Epoch 88/100
469/469 - 52s - loss: 0.1111 - sparse_categorical_accuracy: 0.9642 - val_loss: 0.0996 - val_sparse_c
ategorical_accuracy: 0.9687 - 52s/epoch - 111ms/step
Epoch 89/100
469/469 - 52s - loss: 0.1148 - sparse_categorical_accuracy: 0.9637 - val_loss: 0.0928 - val_sparse_c
ategorical_accuracy: 0.9718 - 52s/epoch - 111ms/step
Epoch 90/100
469/469 - 52s - loss: 0.1108 - sparse_categorical_accuracy: 0.9643 - val_loss: 0.0923 - val_sparse_c
ategorical_accuracy: 0.9708 - 52s/epoch - 111ms/step
Epoch 91/100
469/469 - 52s - loss: 0.1110 - sparse_categorical_accuracy: 0.9653 - val_loss: 0.0942 - val_sparse_c
ategorical_accuracy: 0.9702 - 52s/epoch - 111ms/step
Epoch 92/100
469/469 - 52s - loss: 0.1112 - sparse_categorical_accuracy: 0.9642 - val_loss: 0.0952 - val_sparse_c
ategorical_accuracy: 0.9700 - 52s/epoch - 111ms/step
Epoch 93/100
469/469 - 52s - loss: 0.1091 - sparse_categorical_accuracy: 0.9656 - val_loss: 0.0909 - val_sparse_c
ategorical_accuracy: 0.9715 - 52s/epoch - 111ms/step
Epoch 94/100
469/469 - 52s - loss: 0.1087 - sparse_categorical_accuracy: 0.9652 - val_loss: 0.0970 - val_sparse_c
ategorical_accuracy: 0.9702 - 52s/epoch - 111ms/step
Epoch 95/100
469/469 - 52s - loss: 0.1106 - sparse_categorical_accuracy: 0.9650 - val_loss: 0.0938 - val_sparse_c
ategorical_accuracy: 0.9707 - 52s/epoch - 111ms/step
```

```
Epoch 96/100
469/469 - 52s - loss: 0.1097 - sparse_categorical_accuracy: 0.9647 - val_loss: 0.0944 - val_sparse_c
ategorical_accuracy: 0.9708 - 52s/epoch - 111ms/step
Epoch 97/100
469/469 - 52s - loss: 0.1083 - sparse_categorical_accuracy: 0.9659 - val_loss: 0.1024 - val_sparse_c
ategorical_accuracy: 0.9673 - 52s/epoch - 112ms/step
Epoch 98/100
469/469 - 52s - loss: 0.1076 - sparse_categorical_accuracy: 0.9663 - val_loss: 0.0990 - val_sparse_c
ategorical_accuracy: 0.9682 - 52s/epoch - 112ms/step
Epoch 99/100
469/469 - 52s - loss: 0.1073 - sparse_categorical_accuracy: 0.9660 - val_loss: 0.1059 - val_sparse_c
ategorical_accuracy: 0.9675 - 52s/epoch - 111ms/step
Epoch 100/100
469/469 - 52s - loss: 0.1086 - sparse_categorical_accuracy: 0.9656 - val_loss: 0.0930 - val_sparse_c
ategorical_accuracy: 0.9699 - 52s/epoch - 111ms/step
```

In [12]:

```
model1.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 4s - loss: 0.0930 - sparse_categorical_accuracy: 0.9699 - 4s/epoch - 13ms/step
```

Out[12]:

```
[0.09297644346952438, 0.9699000120162964]
```

Model2: `adam` optimizer with learning rate=e^{-4}, `random_uniform` initializer, `dropout` regularization with rate=0.1.

In [13]:

```
input_layer = Input(shape=(28, 28,1))
x = inception_module(input_layer,
                     filters_1x1=1,
                     filters_3x3_reduce=1,
                     filters_3x3=1,
                     filters_5x5_reduce=1,
                     filters_5x5=1,
                     filters_pool_proj=1,
                     name='inception_3a')
x = tf.keras.layers.BatchNormalization()(x)
x = MaxPool2D((3, 3))(x)
x = tf.keras.layers.Flatten()(x)
x = Dropout(0.1)(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform')(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform')(x)
x = tf.keras.layers.BatchNormalization()(x)
x = Dense(10, activation='softmax', kernel_initializer='random_uniform')(x)
model2 = Model(input_layer, x, name='inception_v1')
```

In [14]:

```
model2.compile(optimizer=tf.keras.optimizers.Adam(0.0001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [15]:

```
history2 = model2.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 53s - loss: 1.9642 - sparse_categorical_accuracy: 0.5036 - val_loss: 1.6752 - val_sparse_c
ategorical_accuracy: 0.7477 - 53s/epoch - 113ms/step
Epoch 2/100
469/469 - 52s - loss: 1.3738 - sparse_categorical_accuracy: 0.7804 - val_loss: 1.0907 - val_sparse_c
ategorical_accuracy: 0.8542 - 52s/epoch - 110ms/step
Epoch 3/100
469/469 - 52s - loss: 0.9748 - sparse_categorical_accuracy: 0.8523 - val_loss: 0.7782 - val_sparse_c
ategorical_accuracy: 0.8883 - 52s/epoch - 110ms/step
Epoch 4/100
469/469 - 52s - loss: 0.7113 - sparse_categorical_accuracy: 0.8871 - val_loss: 0.5566 - val_sparse_c
ategorical_accuracy: 0.9170 - 52s/epoch - 110ms/step
Epoch 5/100
469/469 - 52s - loss: 0.5367 - sparse_categorical_accuracy: 0.9057 - val_loss: 0.4216 - val_sparse_c
ategorical_accuracy: 0.9269 - 52s/epoch - 110ms/step
Epoch 6/100
469/469 - 52s - loss: 0.4213 - sparse_categorical_accuracy: 0.9164 - val_loss: 0.3323 - val_sparse_c
```

```
ategorical_accuracy: 0.9365 - 52s/epoch - 111ms/step
Epoch 7/100
469/469 - 52s - loss: 0.3488 - sparse_categorical_accuracy: 0.9237 - val_loss: 0.2746 - val_sparse_c
ategorical_accuracy: 0.9414 - 52s/epoch - 110ms/step
Epoch 8/100
469/469 - 52s - loss: 0.2983 - sparse_categorical_accuracy: 0.9302 - val_loss: 0.2342 - val_sparse_c
ategorical_accuracy: 0.9449 - 52s/epoch - 110ms/step
Epoch 9/100
469/469 - 52s - loss: 0.2652 - sparse_categorical_accuracy: 0.9328 - val_loss: 0.2077 - val_sparse_c
ategorical_accuracy: 0.9485 - 52s/epoch - 110ms/step
Epoch 10/100
469/469 - 52s - loss: 0.2407 - sparse_categorical_accuracy: 0.9375 - val_loss: 0.1864 - val_sparse_c
ategorical_accuracy: 0.9508 - 52s/epoch - 110ms/step
Epoch 11/100
469/469 - 52s - loss: 0.2215 - sparse_categorical_accuracy: 0.9405 - val_loss: 0.1741 - val_sparse_c
ategorical_accuracy: 0.9532 - 52s/epoch - 110ms/step
Epoch 12/100
469/469 - 52s - loss: 0.2051 - sparse_categorical_accuracy: 0.9432 - val_loss: 0.1630 - val_sparse_c
ategorical_accuracy: 0.9551 - 52s/epoch - 111ms/step
Epoch 13/100
469/469 - 52s - loss: 0.1965 - sparse_categorical_accuracy: 0.9437 - val_loss: 0.1536 - val_sparse_c
ategorical_accuracy: 0.9568 - 52s/epoch - 111ms/step
Epoch 14/100
469/469 - 52s - loss: 0.1867 - sparse_categorical_accuracy: 0.9459 - val_loss: 0.1455 - val_sparse_c
ategorical_accuracy: 0.9587 - 52s/epoch - 111ms/step
Epoch 15/100
469/469 - 52s - loss: 0.1798 - sparse_categorical_accuracy: 0.9480 - val_loss: 0.1395 - val_sparse_c
ategorical_accuracy: 0.9589 - 52s/epoch - 111ms/step
Epoch 16/100
469/469 - 52s - loss: 0.1734 - sparse_categorical_accuracy: 0.9495 - val_loss: 0.1336 - val_sparse_c
ategorical_accuracy: 0.9614 - 52s/epoch - 111ms/step
Epoch 17/100
469/469 - 52s - loss: 0.1661 - sparse_categorical_accuracy: 0.9512 - val_loss: 0.1302 - val_sparse_c
ategorical_accuracy: 0.9628 - 52s/epoch - 111ms/step
Epoch 18/100
469/469 - 52s - loss: 0.1605 - sparse_categorical_accuracy: 0.9526 - val_loss: 0.1274 - val_sparse_c
ategorical_accuracy: 0.9632 - 52s/epoch - 111ms/step
Epoch 19/100
469/469 - 52s - loss: 0.1576 - sparse_categorical_accuracy: 0.9531 - val_loss: 0.1220 - val_sparse_c
ategorical_accuracy: 0.9637 - 52s/epoch - 111ms/step
Epoch 20/100
469/469 - 52s - loss: 0.1543 - sparse_categorical_accuracy: 0.9546 - val_loss: 0.1202 - val_sparse_c
ategorical_accuracy: 0.9640 - 52s/epoch - 111ms/step
Epoch 21/100
469/469 - 52s - loss: 0.1491 - sparse_categorical_accuracy: 0.9556 - val_loss: 0.1189 - val_sparse_c
ategorical_accuracy: 0.9646 - 52s/epoch - 112ms/step
Epoch 22/100
469/469 - 52s - loss: 0.1479 - sparse_categorical_accuracy: 0.9556 - val_loss: 0.1152 - val_sparse_c
ategorical_accuracy: 0.9649 - 52s/epoch - 111ms/step
Epoch 23/100
469/469 - 52s - loss: 0.1475 - sparse_categorical_accuracy: 0.9560 - val_loss: 0.1135 - val_sparse_c
ategorical_accuracy: 0.9654 - 52s/epoch - 112ms/step
Epoch 24/100
469/469 - 52s - loss: 0.1426 - sparse_categorical_accuracy: 0.9572 - val_loss: 0.1128 - val_sparse_c
ategorical_accuracy: 0.9656 - 52s/epoch - 112ms/step
Epoch 25/100
469/469 - 52s - loss: 0.1407 - sparse_categorical_accuracy: 0.9581 - val_loss: 0.1122 - val_sparse_c
ategorical_accuracy: 0.9652 - 52s/epoch - 112ms/step
Epoch 26/100
469/469 - 52s - loss: 0.1397 - sparse_categorical_accuracy: 0.9576 - val_loss: 0.1095 - val_sparse_c
ategorical_accuracy: 0.9666 - 52s/epoch - 112ms/step
Epoch 27/100
469/469 - 52s - loss: 0.1390 - sparse_categorical_accuracy: 0.9581 - val_loss: 0.1089 - val_sparse_c
ategorical_accuracy: 0.9668 - 52s/epoch - 112ms/step
Epoch 28/100
469/469 - 52s - loss: 0.1373 - sparse_categorical_accuracy: 0.9587 - val_loss: 0.1054 - val_sparse_c
ategorical_accuracy: 0.9667 - 52s/epoch - 111ms/step
Epoch 29/100
469/469 - 52s - loss: 0.1333 - sparse_categorical_accuracy: 0.9599 - val_loss: 0.1061 - val_sparse_c
ategorical_accuracy: 0.9672 - 52s/epoch - 112ms/step
Epoch 30/100
469/469 - 52s - loss: 0.1318 - sparse_categorical_accuracy: 0.9607 - val_loss: 0.1037 - val_sparse_c
ategorical_accuracy: 0.9678 - 52s/epoch - 111ms/step
Epoch 31/100
469/469 - 52s - loss: 0.1322 - sparse_categorical_accuracy: 0.9597 - val_loss: 0.1025 - val_sparse_c
ategorical_accuracy: 0.9686 - 52s/epoch - 112ms/step
Epoch 32/100
469/469 - 52s - loss: 0.1306 - sparse_categorical_accuracy: 0.9601 - val_loss: 0.1015 - val_sparse_c
ategorical_accuracy: 0.9691 - 52s/epoch - 112ms/step
Epoch 33/100
469/469 - 52s - loss: 0.1285 - sparse_categorical_accuracy: 0.9607 - val_loss: 0.1008 - val_sparse_c
ategorical_accuracy: 0.9678 - 52s/epoch - 112ms/step
Epoch 34/100
```

```
469/469 - 52s - loss: 0.1270 - sparse_categorical_accuracy: 0.9613 - val_loss: 0.0998 - val_sparse_c
ategorical_accuracy: 0.9681 - 52s/epoch - 112ms/step
Epoch 35/100
469/469 - 52s - loss: 0.1258 - sparse_categorical_accuracy: 0.9613 - val_loss: 0.1008 - val_sparse_c
ategorical_accuracy: 0.9696 - 52s/epoch - 111ms/step
Epoch 36/100
469/469 - 52s - loss: 0.1260 - sparse_categorical_accuracy: 0.9616 - val_loss: 0.0982 - val_sparse_c
ategorical_accuracy: 0.9682 - 52s/epoch - 111ms/step
Epoch 37/100
469/469 - 52s - loss: 0.1259 - sparse_categorical_accuracy: 0.9609 - val_loss: 0.0992 - val_sparse_c
ategorical_accuracy: 0.9678 - 52s/epoch - 111ms/step
Epoch 38/100
469/469 - 52s - loss: 0.1246 - sparse_categorical_accuracy: 0.9619 - val_loss: 0.0965 - val_sparse_c
ategorical_accuracy: 0.9698 - 52s/epoch - 111ms/step
Epoch 39/100
469/469 - 52s - loss: 0.1248 - sparse_categorical_accuracy: 0.9610 - val_loss: 0.0980 - val_sparse_c
ategorical_accuracy: 0.9689 - 52s/epoch - 112ms/step
Epoch 40/100
469/469 - 52s - loss: 0.1217 - sparse_categorical_accuracy: 0.9625 - val_loss: 0.0970 - val_sparse_c
ategorical_accuracy: 0.9676 - 52s/epoch - 111ms/step
Epoch 41/100
469/469 - 52s - loss: 0.1211 - sparse_categorical_accuracy: 0.9627 - val_loss: 0.0954 - val_sparse_c
ategorical_accuracy: 0.9694 - 52s/epoch - 111ms/step
Epoch 42/100
469/469 - 52s - loss: 0.1192 - sparse_categorical_accuracy: 0.9633 - val_loss: 0.0943 - val_sparse_c
ategorical_accuracy: 0.9701 - 52s/epoch - 111ms/step
Epoch 43/100
469/469 - 52s - loss: 0.1192 - sparse_categorical_accuracy: 0.9628 - val_loss: 0.0960 - val_sparse_c
ategorical_accuracy: 0.9686 - 52s/epoch - 111ms/step
Epoch 44/100
469/469 - 52s - loss: 0.1212 - sparse_categorical_accuracy: 0.9625 - val_loss: 0.0940 - val_sparse_c
ategorical_accuracy: 0.9704 - 52s/epoch - 112ms/step
Epoch 45/100
469/469 - 52s - loss: 0.1199 - sparse_categorical_accuracy: 0.9629 - val_loss: 0.0921 - val_sparse_c
ategorical_accuracy: 0.9704 - 52s/epoch - 111ms/step
Epoch 46/100
469/469 - 52s - loss: 0.1170 - sparse_categorical_accuracy: 0.9628 - val_loss: 0.0964 - val_sparse_c
ategorical_accuracy: 0.9691 - 52s/epoch - 111ms/step
Epoch 47/100
469/469 - 52s - loss: 0.1156 - sparse_categorical_accuracy: 0.9639 - val_loss: 0.0924 - val_sparse_c
ategorical_accuracy: 0.9702 - 52s/epoch - 111ms/step
Epoch 48/100
469/469 - 52s - loss: 0.1153 - sparse_categorical_accuracy: 0.9641 - val_loss: 0.0904 - val_sparse_c
ategorical_accuracy: 0.9716 - 52s/epoch - 111ms/step
Epoch 49/100
469/469 - 52s - loss: 0.1159 - sparse_categorical_accuracy: 0.9635 - val_loss: 0.0917 - val_sparse_c
ategorical_accuracy: 0.9699 - 52s/epoch - 111ms/step
Epoch 50/100
469/469 - 52s - loss: 0.1151 - sparse_categorical_accuracy: 0.9642 - val_loss: 0.0909 - val_sparse_c
ategorical_accuracy: 0.9707 - 52s/epoch - 111ms/step
Epoch 51/100
469/469 - 52s - loss: 0.1146 - sparse_categorical_accuracy: 0.9650 - val_loss: 0.0905 - val_sparse_c
ategorical_accuracy: 0.9707 - 52s/epoch - 111ms/step
Epoch 52/100
469/469 - 52s - loss: 0.1130 - sparse_categorical_accuracy: 0.9647 - val_loss: 0.0895 - val_sparse_c
ategorical_accuracy: 0.9714 - 52s/epoch - 111ms/step
Epoch 53/100
469/469 - 52s - loss: 0.1115 - sparse_categorical_accuracy: 0.9653 - val_loss: 0.0887 - val_sparse_c
ategorical_accuracy: 0.9705 - 52s/epoch - 112ms/step
Epoch 54/100
469/469 - 52s - loss: 0.1128 - sparse_categorical_accuracy: 0.9651 - val_loss: 0.0886 - val_sparse_c
ategorical_accuracy: 0.9721 - 52s/epoch - 111ms/step
Epoch 55/100
469/469 - 52s - loss: 0.1129 - sparse_categorical_accuracy: 0.9647 - val_loss: 0.0906 - val_sparse_c
ategorical_accuracy: 0.9701 - 52s/epoch - 112ms/step
Epoch 56/100
469/469 - 52s - loss: 0.1096 - sparse_categorical_accuracy: 0.9654 - val_loss: 0.0879 - val_sparse_c
ategorical_accuracy: 0.9707 - 52s/epoch - 112ms/step
Epoch 57/100
469/469 - 52s - loss: 0.1109 - sparse_categorical_accuracy: 0.9653 - val_loss: 0.0862 - val_sparse_c
ategorical_accuracy: 0.9722 - 52s/epoch - 112ms/step
Epoch 58/100
469/469 - 52s - loss: 0.1100 - sparse_categorical_accuracy: 0.9653 - val_loss: 0.0864 - val_sparse_c
ategorical_accuracy: 0.9713 - 52s/epoch - 111ms/step
Epoch 59/100
469/469 - 52s - loss: 0.1101 - sparse_categorical_accuracy: 0.9653 - val_loss: 0.0873 - val_sparse_c
ategorical_accuracy: 0.9721 - 52s/epoch - 111ms/step
Epoch 60/100
469/469 - 52s - loss: 0.1088 - sparse_categorical_accuracy: 0.9656 - val_loss: 0.0855 - val_sparse_c
ategorical_accuracy: 0.9717 - 52s/epoch - 112ms/step
Epoch 61/100
469/469 - 52s - loss: 0.1074 - sparse_categorical_accuracy: 0.9664 - val_loss: 0.0856 - val_sparse_c
ategorical_accuracy: 0.9717 - 52s/epoch - 112ms/step
```

```
Epoch 62/100
469/469 - 52s - loss: 0.1077 - sparse_categorical_accuracy: 0.9657 - val_loss: 0.0860 - val_sparse_c
ategorical_accuracy: 0.9709 - 52s/epoch - 112ms/step
Epoch 63/100
469/469 - 52s - loss: 0.1091 - sparse_categorical_accuracy: 0.9656 - val_loss: 0.0864 - val_sparse_c
ategorical_accuracy: 0.9725 - 52s/epoch - 112ms/step
Epoch 64/100
469/469 - 52s - loss: 0.1053 - sparse_categorical_accuracy: 0.9670 - val_loss: 0.0853 - val_sparse_c
ategorical_accuracy: 0.9723 - 52s/epoch - 111ms/step
Epoch 65/100
469/469 - 52s - loss: 0.1044 - sparse_categorical_accuracy: 0.9673 - val_loss: 0.0864 - val_sparse_c
ategorical_accuracy: 0.9719 - 52s/epoch - 111ms/step
Epoch 66/100
469/469 - 52s - loss: 0.1059 - sparse_categorical_accuracy: 0.9668 - val_loss: 0.0838 - val_sparse_c
ategorical_accuracy: 0.9725 - 52s/epoch - 111ms/step
Epoch 67/100
469/469 - 52s - loss: 0.1063 - sparse_categorical_accuracy: 0.9666 - val_loss: 0.0845 - val_sparse_c
ategorical_accuracy: 0.9724 - 52s/epoch - 112ms/step
Epoch 68/100
469/469 - 52s - loss: 0.1057 - sparse_categorical_accuracy: 0.9669 - val_loss: 0.0837 - val_sparse_c
ategorical_accuracy: 0.9732 - 52s/epoch - 111ms/step
Epoch 69/100
469/469 - 52s - loss: 0.1050 - sparse_categorical_accuracy: 0.9668 - val_loss: 0.0837 - val_sparse_c
ategorical_accuracy: 0.9722 - 52s/epoch - 112ms/step
Epoch 70/100
469/469 - 52s - loss: 0.1038 - sparse_categorical_accuracy: 0.9674 - val_loss: 0.0829 - val_sparse_c
ategorical_accuracy: 0.9730 - 52s/epoch - 112ms/step
Epoch 71/100
469/469 - 52s - loss: 0.1035 - sparse_categorical_accuracy: 0.9671 - val_loss: 0.0837 - val_sparse_c
ategorical_accuracy: 0.9723 - 52s/epoch - 111ms/step
Epoch 72/100
469/469 - 52s - loss: 0.1012 - sparse_categorical_accuracy: 0.9678 - val_loss: 0.0826 - val_sparse_c
ategorical_accuracy: 0.9739 - 52s/epoch - 111ms/step
Epoch 73/100
469/469 - 52s - loss: 0.1032 - sparse_categorical_accuracy: 0.9675 - val_loss: 0.0823 - val_sparse_c
ategorical_accuracy: 0.9727 - 52s/epoch - 112ms/step
Epoch 74/100
469/469 - 52s - loss: 0.1015 - sparse_categorical_accuracy: 0.9679 - val_loss: 0.0825 - val_sparse_c
ategorical_accuracy: 0.9728 - 52s/epoch - 112ms/step
Epoch 75/100
469/469 - 52s - loss: 0.1016 - sparse_categorical_accuracy: 0.9683 - val_loss: 0.0851 - val_sparse_c
ategorical_accuracy: 0.9711 - 52s/epoch - 111ms/step
Epoch 76/100
469/469 - 52s - loss: 0.1034 - sparse_categorical_accuracy: 0.9668 - val_loss: 0.0819 - val_sparse_c
ategorical_accuracy: 0.9729 - 52s/epoch - 111ms/step
Epoch 77/100
469/469 - 52s - loss: 0.1016 - sparse_categorical_accuracy: 0.9682 - val_loss: 0.0800 - val_sparse_c
ategorical_accuracy: 0.9735 - 52s/epoch - 111ms/step
Epoch 78/100
469/469 - 52s - loss: 0.1010 - sparse_categorical_accuracy: 0.9684 - val_loss: 0.0807 - val_sparse_c
ategorical_accuracy: 0.9735 - 52s/epoch - 112ms/step
Epoch 79/100
469/469 - 52s - loss: 0.1023 - sparse_categorical_accuracy: 0.9675 - val_loss: 0.0832 - val_sparse_c
ategorical_accuracy: 0.9740 - 52s/epoch - 112ms/step
Epoch 80/100
469/469 - 52s - loss: 0.0999 - sparse_categorical_accuracy: 0.9690 - val_loss: 0.0820 - val_sparse_c
ategorical_accuracy: 0.9723 - 52s/epoch - 112ms/step
Epoch 81/100
469/469 - 52s - loss: 0.1010 - sparse_categorical_accuracy: 0.9685 - val_loss: 0.0821 - val_sparse_c
ategorical_accuracy: 0.9729 - 52s/epoch - 112ms/step
Epoch 82/100
469/469 - 52s - loss: 0.1004 - sparse_categorical_accuracy: 0.9687 - val_loss: 0.0805 - val_sparse_c
ategorical_accuracy: 0.9736 - 52s/epoch - 112ms/step
Epoch 83/100
469/469 - 54s - loss: 0.1013 - sparse_categorical_accuracy: 0.9679 - val_loss: 0.0799 - val_sparse_c
ategorical_accuracy: 0.9741 - 54s/epoch - 114ms/step
Epoch 84/100
469/469 - 53s - loss: 0.0986 - sparse_categorical_accuracy: 0.9691 - val_loss: 0.0806 - val_sparse_c
ategorical_accuracy: 0.9729 - 53s/epoch - 112ms/step
Epoch 85/100
469/469 - 52s - loss: 0.0971 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0807 - val_sparse_c
ategorical_accuracy: 0.9733 - 52s/epoch - 112ms/step
Epoch 86/100
469/469 - 53s - loss: 0.0981 - sparse_categorical_accuracy: 0.9690 - val_loss: 0.0809 - val_sparse_c
ategorical_accuracy: 0.9723 - 53s/epoch - 112ms/step
Epoch 87/100
469/469 - 52s - loss: 0.0975 - sparse_categorical_accuracy: 0.9691 - val_loss: 0.0808 - val_sparse_c
ategorical_accuracy: 0.9732 - 52s/epoch - 112ms/step
Epoch 88/100
469/469 - 52s - loss: 0.0968 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0813 - val_sparse_c
ategorical_accuracy: 0.9722 - 52s/epoch - 112ms/step
Epoch 89/100
469/469 - 53s - loss: 0.0969 - sparse_categorical_accuracy: 0.9689 - val_loss: 0.0806 - val_sparse_c
```

```
ategorical_accuracy: 0.9733 - 53s/epoch - 112ms/step
Epoch 90/100
469/469 - 52s - loss: 0.0974 - sparse_categorical_accuracy: 0.9695 - val_loss: 0.0796 - val_sparse_c
ategorical_accuracy: 0.9726 - 52s/epoch - 112ms/step
Epoch 91/100
469/469 - 52s - loss: 0.0975 - sparse_categorical_accuracy: 0.9693 - val_loss: 0.0794 - val_sparse_c
ategorical_accuracy: 0.9743 - 52s/epoch - 111ms/step
Epoch 92/100
469/469 - 52s - loss: 0.0961 - sparse_categorical_accuracy: 0.9695 - val_loss: 0.0789 - val_sparse_c
ategorical_accuracy: 0.9740 - 52s/epoch - 112ms/step
Epoch 93/100
469/469 - 52s - loss: 0.0991 - sparse_categorical_accuracy: 0.9689 - val_loss: 0.0792 - val_sparse_c
ategorical_accuracy: 0.9740 - 52s/epoch - 111ms/step
Epoch 94/100
469/469 - 52s - loss: 0.0976 - sparse_categorical_accuracy: 0.9689 - val_loss: 0.0794 - val_sparse_c
ategorical_accuracy: 0.9736 - 52s/epoch - 111ms/step
Epoch 95/100
469/469 - 52s - loss: 0.0966 - sparse_categorical_accuracy: 0.9700 - val_loss: 0.0783 - val_sparse_c
ategorical_accuracy: 0.9737 - 52s/epoch - 112ms/step
Epoch 96/100
469/469 - 52s - loss: 0.0964 - sparse_categorical_accuracy: 0.9694 - val_loss: 0.0779 - val_sparse_c
ategorical_accuracy: 0.9740 - 52s/epoch - 112ms/step
Epoch 97/100
469/469 - 52s - loss: 0.0947 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0800 - val_sparse_c
ategorical_accuracy: 0.9744 - 52s/epoch - 112ms/step
Epoch 98/100
469/469 - 52s - loss: 0.0962 - sparse_categorical_accuracy: 0.9690 - val_loss: 0.0799 - val_sparse_c
ategorical_accuracy: 0.9736 - 52s/epoch - 111ms/step
Epoch 99/100
469/469 - 52s - loss: 0.0964 - sparse_categorical_accuracy: 0.9701 - val_loss: 0.0786 - val_sparse_c
ategorical_accuracy: 0.9739 - 52s/epoch - 111ms/step
Epoch 100/100
469/469 - 52s - loss: 0.0965 - sparse_categorical_accuracy: 0.9695 - val_loss: 0.0780 - val_sparse_c
ategorical_accuracy: 0.9742 - 52s/epoch - 111ms/step
```

In [16]:

```
model2.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 4s - loss: 0.0780 - sparse_categorical_accuracy: 0.9742 - 4s/epoch - 14ms/step
```

Out[16]:

```
[0.07802172750234604, 0.9742000102996826]
```

Model3: `adam` optimizer with learning rate=e^{-3}, `random_normal` initializer, `dropout` regularization with rate=0.1.

In [17]:

```
input_layer = Input(shape=(28, 28,1))
x = inception_module(input_layer,
                     filters_1x1=1,
                     filters_3x3_reduce=1,
                     filters_3x3=1,
                     filters_5x5_reduce=1,
                     filters_5x5=1,
                     filters_pool_proj=1,
                     name='inception_3a')
x = tf.keras.layers.BatchNormalization()(x)
x = MaxPool2D((3, 3))(x)
x = tf.keras.layers.Flatten()(x)
x = Dropout(0.1)(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal')(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal')(x)
x = tf.keras.layers.BatchNormalization()(x)
x = Dense(10, activation='softmax', kernel_initializer='random_uniform')(x)
model3 = Model(input_layer, x, name='inception_v1')
```

In [18]:

```
model3.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [19]:

```
history3 = model3.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 54s - loss: 0.9932 - sparse_categorical_accuracy: 0.7682 - val_loss: 0.3869 - val_sparse_c
ategorical_accuracy: 0.8969 - 54s/epoch - 114ms/step
Epoch 2/100
469/469 - 52s - loss: 0.4003 - sparse_categorical_accuracy: 0.8785 - val_loss: 0.2898 - val_sparse_c
ategorical_accuracy: 0.9084 - 52s/epoch - 111ms/step
Epoch 3/100
469/469 - 52s - loss: 0.3441 - sparse_categorical_accuracy: 0.8905 - val_loss: 0.2455 - val_sparse_c
ategorical_accuracy: 0.9250 - 52s/epoch - 111ms/step
Epoch 4/100
469/469 - 52s - loss: 0.3152 - sparse_categorical_accuracy: 0.8994 - val_loss: 0.2297 - val_sparse_c
ategorical_accuracy: 0.9277 - 52s/epoch - 111ms/step
Epoch 5/100
469/469 - 52s - loss: 0.2865 - sparse_categorical_accuracy: 0.9082 - val_loss: 0.2129 - val_sparse_c
ategorical_accuracy: 0.9337 - 52s/epoch - 111ms/step
Epoch 6/100
469/469 - 52s - loss: 0.2523 - sparse_categorical_accuracy: 0.9194 - val_loss: 0.1935 - val_sparse_c
ategorical_accuracy: 0.9378 - 52s/epoch - 111ms/step
Epoch 7/100
469/469 - 52s - loss: 0.2287 - sparse_categorical_accuracy: 0.9278 - val_loss: 0.1886 - val_sparse_c
ategorical_accuracy: 0.9423 - 52s/epoch - 111ms/step
Epoch 8/100
469/469 - 52s - loss: 0.2146 - sparse_categorical_accuracy: 0.9320 - val_loss: 0.1999 - val_sparse_c
ategorical_accuracy: 0.9385 - 52s/epoch - 111ms/step
Epoch 9/100
469/469 - 52s - loss: 0.2040 - sparse_categorical_accuracy: 0.9365 - val_loss: 0.1516 - val_sparse_c
ategorical_accuracy: 0.9544 - 52s/epoch - 111ms/step
Epoch 10/100
469/469 - 52s - loss: 0.1950 - sparse_categorical_accuracy: 0.9380 - val_loss: 0.1567 - val_sparse_c
ategorical_accuracy: 0.9533 - 52s/epoch - 111ms/step
Epoch 11/100
469/469 - 52s - loss: 0.1867 - sparse_categorical_accuracy: 0.9409 - val_loss: 0.1409 - val_sparse_c
ategorical_accuracy: 0.9563 - 52s/epoch - 111ms/step
Epoch 12/100
469/469 - 52s - loss: 0.1820 - sparse_categorical_accuracy: 0.9431 - val_loss: 0.1387 - val_sparse_c
ategorical_accuracy: 0.9576 - 52s/epoch - 112ms/step
Epoch 13/100
469/469 - 53s - loss: 0.1767 - sparse_categorical_accuracy: 0.9443 - val_loss: 0.1388 - val_sparse_c
ategorical_accuracy: 0.9583 - 53s/epoch - 112ms/step
Epoch 14/100
469/469 - 53s - loss: 0.1712 - sparse_categorical_accuracy: 0.9459 - val_loss: 0.1431 - val_sparse_c
ategorical_accuracy: 0.9559 - 53s/epoch - 112ms/step
Epoch 15/100
469/469 - 52s - loss: 0.1681 - sparse_categorical_accuracy: 0.9463 - val_loss: 0.1291 - val_sparse_c
ategorical_accuracy: 0.9611 - 52s/epoch - 112ms/step
Epoch 16/100
469/469 - 52s - loss: 0.1655 - sparse_categorical_accuracy: 0.9474 - val_loss: 0.1298 - val_sparse_c
ategorical_accuracy: 0.9610 - 52s/epoch - 112ms/step
Epoch 17/100
469/469 - 52s - loss: 0.1605 - sparse_categorical_accuracy: 0.9501 - val_loss: 0.1327 - val_sparse_c
ategorical_accuracy: 0.9585 - 52s/epoch - 112ms/step
Epoch 18/100
469/469 - 53s - loss: 0.1576 - sparse_categorical_accuracy: 0.9499 - val_loss: 0.1264 - val_sparse_c
ategorical_accuracy: 0.9615 - 53s/epoch - 112ms/step
Epoch 19/100
469/469 - 53s - loss: 0.1592 - sparse_categorical_accuracy: 0.9495 - val_loss: 0.1289 - val_sparse_c
ategorical_accuracy: 0.9595 - 53s/epoch - 112ms/step
Epoch 20/100
469/469 - 53s - loss: 0.1539 - sparse_categorical_accuracy: 0.9517 - val_loss: 0.1262 - val_sparse_c
ategorical_accuracy: 0.9620 - 53s/epoch - 112ms/step
Epoch 21/100
469/469 - 52s - loss: 0.1511 - sparse_categorical_accuracy: 0.9521 - val_loss: 0.1323 - val_sparse_c
ategorical_accuracy: 0.9593 - 52s/epoch - 112ms/step
Epoch 22/100
469/469 - 52s - loss: 0.1501 - sparse_categorical_accuracy: 0.9521 - val_loss: 0.1335 - val_sparse_c
ategorical_accuracy: 0.9589 - 52s/epoch - 112ms/step
Epoch 23/100
469/469 - 52s - loss: 0.1502 - sparse_categorical_accuracy: 0.9519 - val_loss: 0.1189 - val_sparse_c
ategorical_accuracy: 0.9644 - 52s/epoch - 112ms/step
Epoch 24/100
469/469 - 52s - loss: 0.1470 - sparse_categorical_accuracy: 0.9532 - val_loss: 0.1155 - val_sparse_c
ategorical_accuracy: 0.9657 - 52s/epoch - 112ms/step
Epoch 25/100
469/469 - 52s - loss: 0.1468 - sparse_categorical_accuracy: 0.9528 - val_loss: 0.1266 - val_sparse_c
```

```
ategorical_accuracy: 0.9626 - 52s/epoch - 112ms/step
Epoch 26/100
469/469 - 52s - loss: 0.1414 - sparse_categorical_accuracy: 0.9546 - val_loss: 0.1147 - val_sparse_c
ategorical_accuracy: 0.9654 - 52s/epoch - 111ms/step
Epoch 27/100
469/469 - 52s - loss: 0.1437 - sparse_categorical_accuracy: 0.9538 - val_loss: 0.1147 - val_sparse_c
ategorical_accuracy: 0.9654 - 52s/epoch - 111ms/step
Epoch 28/100
469/469 - 52s - loss: 0.1415 - sparse_categorical_accuracy: 0.9545 - val_loss: 0.1339 - val_sparse_c
ategorical_accuracy: 0.9590 - 52s/epoch - 111ms/step
Epoch 29/100
469/469 - 52s - loss: 0.1369 - sparse_categorical_accuracy: 0.9565 - val_loss: 0.1206 - val_sparse_c
ategorical_accuracy: 0.9632 - 52s/epoch - 112ms/step
Epoch 30/100
469/469 - 52s - loss: 0.1398 - sparse_categorical_accuracy: 0.9555 - val_loss: 0.1100 - val_sparse_c
ategorical_accuracy: 0.9658 - 52s/epoch - 112ms/step
Epoch 31/100
469/469 - 52s - loss: 0.1383 - sparse_categorical_accuracy: 0.9558 - val_loss: 0.1165 - val_sparse_c
ategorical_accuracy: 0.9649 - 52s/epoch - 112ms/step
Epoch 32/100
469/469 - 52s - loss: 0.1386 - sparse_categorical_accuracy: 0.9557 - val_loss: 0.1019 - val_sparse_c
ategorical_accuracy: 0.9704 - 52s/epoch - 112ms/step
Epoch 33/100
469/469 - 52s - loss: 0.1351 - sparse_categorical_accuracy: 0.9570 - val_loss: 0.1215 - val_sparse_c
ategorical_accuracy: 0.9612 - 52s/epoch - 112ms/step
Epoch 34/100
469/469 - 52s - loss: 0.1367 - sparse_categorical_accuracy: 0.9565 - val_loss: 0.1020 - val_sparse_c
ategorical_accuracy: 0.9688 - 52s/epoch - 112ms/step
Epoch 35/100
469/469 - 53s - loss: 0.1352 - sparse_categorical_accuracy: 0.9562 - val_loss: 0.1136 - val_sparse_c
ategorical_accuracy: 0.9656 - 53s/epoch - 112ms/step
Epoch 36/100
469/469 - 52s - loss: 0.1328 - sparse_categorical_accuracy: 0.9579 - val_loss: 0.1111 - val_sparse_c
ategorical_accuracy: 0.9668 - 52s/epoch - 111ms/step
Epoch 37/100
469/469 - 53s - loss: 0.1334 - sparse_categorical_accuracy: 0.9577 - val_loss: 0.1075 - val_sparse_c
ategorical_accuracy: 0.9670 - 53s/epoch - 112ms/step
Epoch 38/100
469/469 - 52s - loss: 0.1325 - sparse_categorical_accuracy: 0.9577 - val_loss: 0.1188 - val_sparse_c
ategorical_accuracy: 0.9612 - 52s/epoch - 112ms/step
Epoch 39/100
469/469 - 52s - loss: 0.1328 - sparse_categorical_accuracy: 0.9580 - val_loss: 0.1054 - val_sparse_c
ategorical_accuracy: 0.9676 - 52s/epoch - 112ms/step
Epoch 40/100
469/469 - 52s - loss: 0.1286 - sparse_categorical_accuracy: 0.9592 - val_loss: 0.0996 - val_sparse_c
ategorical_accuracy: 0.9698 - 52s/epoch - 112ms/step
Epoch 41/100
469/469 - 53s - loss: 0.1279 - sparse_categorical_accuracy: 0.9592 - val_loss: 0.1005 - val_sparse_c
ategorical_accuracy: 0.9688 - 53s/epoch - 112ms/step
Epoch 42/100
469/469 - 52s - loss: 0.1281 - sparse_categorical_accuracy: 0.9589 - val_loss: 0.1032 - val_sparse_c
ategorical_accuracy: 0.9669 - 52s/epoch - 112ms/step
Epoch 43/100
469/469 - 52s - loss: 0.1260 - sparse_categorical_accuracy: 0.9597 - val_loss: 0.1080 - val_sparse_c
ategorical_accuracy: 0.9662 - 52s/epoch - 112ms/step
Epoch 44/100
469/469 - 52s - loss: 0.1309 - sparse_categorical_accuracy: 0.9579 - val_loss: 0.0987 - val_sparse_c
ategorical_accuracy: 0.9703 - 52s/epoch - 111ms/step
Epoch 45/100
469/469 - 52s - loss: 0.1259 - sparse_categorical_accuracy: 0.9601 - val_loss: 0.1079 - val_sparse_c
ategorical_accuracy: 0.9662 - 52s/epoch - 111ms/step
Epoch 46/100
469/469 - 52s - loss: 0.1276 - sparse_categorical_accuracy: 0.9593 - val_loss: 0.1204 - val_sparse_c
ategorical_accuracy: 0.9624 - 52s/epoch - 112ms/step
Epoch 47/100
469/469 - 52s - loss: 0.1254 - sparse_categorical_accuracy: 0.9600 - val_loss: 0.1013 - val_sparse_c
ategorical_accuracy: 0.9688 - 52s/epoch - 112ms/step
Epoch 48/100
469/469 - 52s - loss: 0.1240 - sparse_categorical_accuracy: 0.9609 - val_loss: 0.0982 - val_sparse_c
ategorical_accuracy: 0.9695 - 52s/epoch - 112ms/step
Epoch 49/100
469/469 - 52s - loss: 0.1256 - sparse_categorical_accuracy: 0.9597 - val_loss: 0.1005 - val_sparse_c
ategorical_accuracy: 0.9694 - 52s/epoch - 112ms/step
Epoch 50/100
469/469 - 52s - loss: 0.1250 - sparse_categorical_accuracy: 0.9600 - val_loss: 0.0960 - val_sparse_c
ategorical_accuracy: 0.9704 - 52s/epoch - 112ms/step
Epoch 51/100
469/469 - 52s - loss: 0.1225 - sparse_categorical_accuracy: 0.9607 - val_loss: 0.1048 - val_sparse_c
ategorical_accuracy: 0.9679 - 52s/epoch - 112ms/step
Epoch 52/100
469/469 - 52s - loss: 0.1216 - sparse_categorical_accuracy: 0.9613 - val_loss: 0.1195 - val_sparse_c
ategorical_accuracy: 0.9623 - 52s/epoch - 112ms/step
Epoch 53/100
```

```
469/469 - 53s - loss: 0.1210 - sparse_categorical_accuracy: 0.9612 - val_loss: 0.1014 - val_sparse_c
ategorical_accuracy: 0.9676 - 53s/epoch - 112ms/step
Epoch 54/100
469/469 - 52s - loss: 0.1232 - sparse_categorical_accuracy: 0.9598 - val_loss: 0.0963 - val_sparse_c
ategorical_accuracy: 0.9697 - 52s/epoch - 112ms/step
Epoch 55/100
469/469 - 52s - loss: 0.1201 - sparse_categorical_accuracy: 0.9620 - val_loss: 0.0959 - val_sparse_c
ategorical_accuracy: 0.9710 - 52s/epoch - 112ms/step
Epoch 56/100
469/469 - 52s - loss: 0.1219 - sparse_categorical_accuracy: 0.9613 - val_loss: 0.0977 - val_sparse_c
ategorical_accuracy: 0.9716 - 52s/epoch - 111ms/step
Epoch 57/100
469/469 - 52s - loss: 0.1190 - sparse_categorical_accuracy: 0.9611 - val_loss: 0.0917 - val_sparse_c
ategorical_accuracy: 0.9730 - 52s/epoch - 112ms/step
Epoch 58/100
469/469 - 53s - loss: 0.1185 - sparse_categorical_accuracy: 0.9616 - val_loss: 0.1123 - val_sparse_c
ategorical_accuracy: 0.9661 - 53s/epoch - 112ms/step
Epoch 59/100
469/469 - 52s - loss: 0.1194 - sparse_categorical_accuracy: 0.9618 - val_loss: 0.1012 - val_sparse_c
ategorical_accuracy: 0.9684 - 52s/epoch - 112ms/step
Epoch 60/100
469/469 - 52s - loss: 0.1179 - sparse_categorical_accuracy: 0.9620 - val_loss: 0.0911 - val_sparse_c
ategorical_accuracy: 0.9717 - 52s/epoch - 112ms/step
Epoch 61/100
469/469 - 52s - loss: 0.1170 - sparse_categorical_accuracy: 0.9625 - val_loss: 0.0981 - val_sparse_c
ategorical_accuracy: 0.9689 - 52s/epoch - 112ms/step
Epoch 62/100
469/469 - 52s - loss: 0.1152 - sparse_categorical_accuracy: 0.9629 - val_loss: 0.0913 - val_sparse_c
ategorical_accuracy: 0.9720 - 52s/epoch - 111ms/step
Epoch 63/100
469/469 - 52s - loss: 0.1149 - sparse_categorical_accuracy: 0.9635 - val_loss: 0.1042 - val_sparse_c
ategorical_accuracy: 0.9692 - 52s/epoch - 112ms/step
Epoch 64/100
469/469 - 53s - loss: 0.1155 - sparse_categorical_accuracy: 0.9635 - val_loss: 0.1025 - val_sparse_c
ategorical_accuracy: 0.9696 - 53s/epoch - 112ms/step
Epoch 65/100
469/469 - 52s - loss: 0.1164 - sparse_categorical_accuracy: 0.9628 - val_loss: 0.1004 - val_sparse_c
ategorical_accuracy: 0.9700 - 52s/epoch - 111ms/step
Epoch 66/100
469/469 - 52s - loss: 0.1119 - sparse_categorical_accuracy: 0.9637 - val_loss: 0.0980 - val_sparse_c
ategorical_accuracy: 0.9703 - 52s/epoch - 111ms/step
Epoch 67/100
469/469 - 52s - loss: 0.1126 - sparse_categorical_accuracy: 0.9632 - val_loss: 0.0923 - val_sparse_c
ategorical_accuracy: 0.9711 - 52s/epoch - 112ms/step
Epoch 68/100
469/469 - 52s - loss: 0.1119 - sparse_categorical_accuracy: 0.9642 - val_loss: 0.0884 - val_sparse_c
ategorical_accuracy: 0.9730 - 52s/epoch - 111ms/step
Epoch 69/100
469/469 - 52s - loss: 0.1079 - sparse_categorical_accuracy: 0.9653 - val_loss: 0.0913 - val_sparse_c
ategorical_accuracy: 0.9727 - 52s/epoch - 112ms/step
Epoch 70/100
469/469 - 52s - loss: 0.1087 - sparse_categorical_accuracy: 0.9654 - val_loss: 0.0891 - val_sparse_c
ategorical_accuracy: 0.9739 - 52s/epoch - 112ms/step
Epoch 71/100
469/469 - 53s - loss: 0.1083 - sparse_categorical_accuracy: 0.9654 - val_loss: 0.0832 - val_sparse_c
ategorical_accuracy: 0.9756 - 53s/epoch - 112ms/step
Epoch 72/100
469/469 - 52s - loss: 0.1043 - sparse_categorical_accuracy: 0.9662 - val_loss: 0.1004 - val_sparse_c
ategorical_accuracy: 0.9687 - 52s/epoch - 111ms/step
Epoch 73/100
469/469 - 52s - loss: 0.1045 - sparse_categorical_accuracy: 0.9662 - val_loss: 0.0865 - val_sparse_c
ategorical_accuracy: 0.9738 - 52s/epoch - 111ms/step
Epoch 74/100
469/469 - 52s - loss: 0.1047 - sparse_categorical_accuracy: 0.9665 - val_loss: 0.0930 - val_sparse_c
ategorical_accuracy: 0.9713 - 52s/epoch - 111ms/step
Epoch 75/100
469/469 - 52s - loss: 0.1052 - sparse_categorical_accuracy: 0.9660 - val_loss: 0.0933 - val_sparse_c
ategorical_accuracy: 0.9711 - 52s/epoch - 112ms/step
Epoch 76/100
469/469 - 52s - loss: 0.1021 - sparse_categorical_accuracy: 0.9678 - val_loss: 0.0816 - val_sparse_c
ategorical_accuracy: 0.9771 - 52s/epoch - 111ms/step
Epoch 77/100
469/469 - 52s - loss: 0.1025 - sparse_categorical_accuracy: 0.9668 - val_loss: 0.0974 - val_sparse_c
ategorical_accuracy: 0.9695 - 52s/epoch - 112ms/step
Epoch 78/100
469/469 - 52s - loss: 0.1040 - sparse_categorical_accuracy: 0.9664 - val_loss: 0.0810 - val_sparse_c
ategorical_accuracy: 0.9759 - 52s/epoch - 112ms/step
Epoch 79/100
469/469 - 52s - loss: 0.0991 - sparse_categorical_accuracy: 0.9679 - val_loss: 0.0807 - val_sparse_c
ategorical_accuracy: 0.9744 - 52s/epoch - 112ms/step
Epoch 80/100
469/469 - 52s - loss: 0.1009 - sparse_categorical_accuracy: 0.9678 - val_loss: 0.0836 - val_sparse_c
ategorical_accuracy: 0.9740 - 52s/epoch - 111ms/step
```

```
Epoch 81/100
469/469 - 53s - loss: 0.0993 - sparse_categorical_accuracy: 0.9683 - val_loss: 0.0875 - val_sparse_c
ategorical_accuracy: 0.9728 - 53s/epoch - 112ms/step
Epoch 82/100
469/469 - 52s - loss: 0.0984 - sparse_categorical_accuracy: 0.9682 - val_loss: 0.0822 - val_sparse_c
ategorical_accuracy: 0.9742 - 52s/epoch - 111ms/step
Epoch 83/100
469/469 - 52s - loss: 0.0973 - sparse_categorical_accuracy: 0.9690 - val_loss: 0.0832 - val_sparse_c
ategorical_accuracy: 0.9749 - 52s/epoch - 112ms/step
Epoch 84/100
469/469 - 52s - loss: 0.0977 - sparse_categorical_accuracy: 0.9688 - val_loss: 0.0889 - val_sparse_c
ategorical_accuracy: 0.9724 - 52s/epoch - 111ms/step
Epoch 85/100
469/469 - 52s - loss: 0.0981 - sparse_categorical_accuracy: 0.9690 - val_loss: 0.0805 - val_sparse_c
ategorical_accuracy: 0.9743 - 52s/epoch - 112ms/step
Epoch 86/100
469/469 - 52s - loss: 0.0986 - sparse_categorical_accuracy: 0.9691 - val_loss: 0.0850 - val_sparse_c
ategorical_accuracy: 0.9746 - 52s/epoch - 112ms/step
Epoch 87/100
469/469 - 52s - loss: 0.0981 - sparse_categorical_accuracy: 0.9682 - val_loss: 0.0776 - val_sparse_c
ategorical_accuracy: 0.9763 - 52s/epoch - 112ms/step
Epoch 88/100
469/469 - 52s - loss: 0.0955 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0806 - val_sparse_c
ategorical_accuracy: 0.9765 - 52s/epoch - 111ms/step
Epoch 89/100
469/469 - 52s - loss: 0.0955 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0826 - val_sparse_c
ategorical_accuracy: 0.9738 - 52s/epoch - 112ms/step
Epoch 90/100
469/469 - 52s - loss: 0.0958 - sparse_categorical_accuracy: 0.9693 - val_loss: 0.0792 - val_sparse_c
ategorical_accuracy: 0.9765 - 52s/epoch - 111ms/step
Epoch 91/100
469/469 - 52s - loss: 0.0957 - sparse_categorical_accuracy: 0.9691 - val_loss: 0.0874 - val_sparse_c
ategorical_accuracy: 0.9732 - 52s/epoch - 111ms/step
Epoch 92/100
469/469 - 53s - loss: 0.0951 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0842 - val_sparse_c
ategorical_accuracy: 0.9739 - 53s/epoch - 112ms/step
Epoch 93/100
469/469 - 52s - loss: 0.0929 - sparse_categorical_accuracy: 0.9696 - val_loss: 0.0834 - val_sparse_c
ategorical_accuracy: 0.9750 - 52s/epoch - 112ms/step
Epoch 94/100
469/469 - 52s - loss: 0.0937 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0840 - val_sparse_c
ategorical_accuracy: 0.9750 - 52s/epoch - 112ms/step
Epoch 95/100
469/469 - 52s - loss: 0.0930 - sparse_categorical_accuracy: 0.9704 - val_loss: 0.0827 - val_sparse_c
ategorical_accuracy: 0.9747 - 52s/epoch - 112ms/step
Epoch 96/100
469/469 - 52s - loss: 0.0972 - sparse_categorical_accuracy: 0.9689 - val_loss: 0.0783 - val_sparse_c
ategorical_accuracy: 0.9759 - 52s/epoch - 112ms/step
Epoch 97/100
469/469 - 52s - loss: 0.0928 - sparse_categorical_accuracy: 0.9703 - val_loss: 0.0823 - val_sparse_c
ategorical_accuracy: 0.9746 - 52s/epoch - 112ms/step
Epoch 98/100
469/469 - 53s - loss: 0.0925 - sparse_categorical_accuracy: 0.9702 - val_loss: 0.0815 - val_sparse_c
ategorical_accuracy: 0.9750 - 53s/epoch - 112ms/step
Epoch 99/100
469/469 - 52s - loss: 0.0929 - sparse_categorical_accuracy: 0.9706 - val_loss: 0.0796 - val_sparse_c
ategorical_accuracy: 0.9753 - 52s/epoch - 112ms/step
Epoch 100/100
469/469 - 52s - loss: 0.0937 - sparse_categorical_accuracy: 0.9695 - val_loss: 0.0802 - val_sparse_c
ategorical_accuracy: 0.9749 - 52s/epoch - 112ms/step
```

In [20]:

```
model3.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 4s - loss: 0.0802 - sparse_categorical_accuracy: 0.9749 - 4s/epoch - 13ms/step
```

Out[20]:

```
[0.0802013948559761, 0.9749000072479248]
```

Model4: `adam` optimizer with learning rate=e^{-4}, `random_normal` initializer, `dropout` regularization with rate=0.1.

In [21]:

```
input_layer = Input(shape=(28, 28,1))
x = inception_module(input_layer,
                     filters_1x1=1,
                     filters_3x3_reduce=1,
                     filters_3x3=1,
                     filters_5x5_reduce=1,
                     filters_5x5=1,
                     filters_pool_proj=1,
                     name='inception_3a')
x = tf.keras.layers.BatchNormalization()(x)
x = MaxPool2D((3, 3))(x)
x = tf.keras.layers.Flatten()(x)
x = Dropout(0.1)(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal')(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal')(x)
x = tf.keras.layers.BatchNormalization()(x)
x = Dense(10, activation='softmax', kernel_initializer='random_uniform')(x)
model4 = Model(input_layer, x, name='inception_v1')
```

In [22]:

```
model4.compile(optimizer=tf.keras.optimizers.Adam(0.0001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [23]:

```
history4 = model4.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 54s - loss: 2.0564 - sparse_categorical_accuracy: 0.4430 - val_loss: 1.7816 - val_sparse_c
ategorical_accuracy: 0.6562 - 54s/epoch - 114ms/step
Epoch 2/100
469/469 - 52s - loss: 1.5660 - sparse_categorical_accuracy: 0.7212 - val_loss: 1.2929 - val_sparse_c
ategorical_accuracy: 0.8171 - 52s/epoch - 112ms/step
Epoch 3/100
469/469 - 52s - loss: 1.1771 - sparse_categorical_accuracy: 0.8075 - val_loss: 0.9574 - val_sparse_c
ategorical_accuracy: 0.8673 - 52s/epoch - 112ms/step
Epoch 4/100
469/469 - 52s - loss: 0.8664 - sparse_categorical_accuracy: 0.8464 - val_loss: 0.6902 - val_sparse_c
ategorical_accuracy: 0.8890 - 52s/epoch - 112ms/step
Epoch 5/100
469/469 - 52s - loss: 0.6703 - sparse_categorical_accuracy: 0.8669 - val_loss: 0.5344 - val_sparse_c
ategorical_accuracy: 0.8998 - 52s/epoch - 111ms/step
Epoch 6/100
469/469 - 52s - loss: 0.5501 - sparse_categorical_accuracy: 0.8763 - val_loss: 0.4366 - val_sparse_c
ategorical_accuracy: 0.9092 - 52s/epoch - 112ms/step
Epoch 7/100
469/469 - 53s - loss: 0.4694 - sparse_categorical_accuracy: 0.8859 - val_loss: 0.3730 - val_sparse_c
ategorical_accuracy: 0.9162 - 53s/epoch - 112ms/step
Epoch 8/100
469/469 - 53s - loss: 0.4172 - sparse_categorical_accuracy: 0.8914 - val_loss: 0.3310 - val_sparse_c
ategorical_accuracy: 0.9208 - 53s/epoch - 112ms/step
Epoch 9/100
469/469 - 53s - loss: 0.3797 - sparse_categorical_accuracy: 0.8967 - val_loss: 0.2967 - val_sparse_c
ategorical_accuracy: 0.9236 - 53s/epoch - 112ms/step
Epoch 10/100
469/469 - 53s - loss: 0.3518 - sparse_categorical_accuracy: 0.9003 - val_loss: 0.2710 - val_sparse_c
ategorical_accuracy: 0.9289 - 53s/epoch - 112ms/step
Epoch 11/100
469/469 - 53s - loss: 0.3318 - sparse_categorical_accuracy: 0.9034 - val_loss: 0.2570 - val_sparse_c
ategorical_accuracy: 0.9279 - 53s/epoch - 113ms/step
Epoch 12/100
469/469 - 53s - loss: 0.3140 - sparse_categorical_accuracy: 0.9064 - val_loss: 0.2379 - val_sparse_c
ategorical_accuracy: 0.9337 - 53s/epoch - 112ms/step
Epoch 13/100
469/469 - 53s - loss: 0.3035 - sparse_categorical_accuracy: 0.9081 - val_loss: 0.2283 - val_sparse_c
ategorical_accuracy: 0.9350 - 53s/epoch - 112ms/step
Epoch 14/100
469/469 - 53s - loss: 0.2901 - sparse_categorical_accuracy: 0.9120 - val_loss: 0.2190 - val_sparse_c
ategorical_accuracy: 0.9363 - 53s/epoch - 112ms/step
Epoch 15/100
469/469 - 53s - loss: 0.2818 - sparse_categorical_accuracy: 0.9152 - val_loss: 0.2114 - val_sparse_c
ategorical_accuracy: 0.9384 - 53s/epoch - 113ms/step
Epoch 16/100
```

```
469/469 - 53s - loss: 0.2746 - sparse_categorical_accuracy: 0.9158 - val_loss: 0.2031 - val_sparse_c
ategorical_accuracy: 0.9400 - 53s/epoch - 113ms/step
Epoch 17/100
469/469 - 53s - loss: 0.2683 - sparse_categorical_accuracy: 0.9164 - val_loss: 0.1997 - val_sparse_c
ategorical_accuracy: 0.9398 - 53s/epoch - 112ms/step
Epoch 18/100
469/469 - 53s - loss: 0.2610 - sparse_categorical_accuracy: 0.9181 - val_loss: 0.1923 - val_sparse_c
ategorical_accuracy: 0.9421 - 53s/epoch - 113ms/step
Epoch 19/100
469/469 - 53s - loss: 0.2549 - sparse_categorical_accuracy: 0.9204 - val_loss: 0.1878 - val_sparse_c
ategorical_accuracy: 0.9423 - 53s/epoch - 112ms/step
Epoch 20/100
469/469 - 52s - loss: 0.2504 - sparse_categorical_accuracy: 0.9215 - val_loss: 0.1822 - val_sparse_c
ategorical_accuracy: 0.9440 - 52s/epoch - 112ms/step
Epoch 21/100
469/469 - 53s - loss: 0.2444 - sparse_categorical_accuracy: 0.9238 - val_loss: 0.1817 - val_sparse_c
ategorical_accuracy: 0.9441 - 53s/epoch - 113ms/step
Epoch 22/100
469/469 - 53s - loss: 0.2409 - sparse_categorical_accuracy: 0.9233 - val_loss: 0.1769 - val_sparse_c
ategorical_accuracy: 0.9454 - 53s/epoch - 112ms/step
Epoch 23/100
469/469 - 52s - loss: 0.2377 - sparse_categorical_accuracy: 0.9254 - val_loss: 0.1773 - val_sparse_c
ategorical_accuracy: 0.9455 - 52s/epoch - 112ms/step
Epoch 24/100
469/469 - 53s - loss: 0.2356 - sparse_categorical_accuracy: 0.9249 - val_loss: 0.1710 - val_sparse_c
ategorical_accuracy: 0.9470 - 53s/epoch - 112ms/step
Epoch 25/100
469/469 - 53s - loss: 0.2306 - sparse_categorical_accuracy: 0.9264 - val_loss: 0.1687 - val_sparse_c
ategorical_accuracy: 0.9482 - 53s/epoch - 112ms/step
Epoch 26/100
469/469 - 53s - loss: 0.2300 - sparse_categorical_accuracy: 0.9263 - val_loss: 0.1636 - val_sparse_c
ategorical_accuracy: 0.9499 - 53s/epoch - 112ms/step
Epoch 27/100
469/469 - 53s - loss: 0.2252 - sparse_categorical_accuracy: 0.9281 - val_loss: 0.1622 - val_sparse_c
ategorical_accuracy: 0.9499 - 53s/epoch - 113ms/step
Epoch 28/100
469/469 - 53s - loss: 0.2200 - sparse_categorical_accuracy: 0.9301 - val_loss: 0.1631 - val_sparse_c
ategorical_accuracy: 0.9493 - 53s/epoch - 112ms/step
Epoch 29/100
469/469 - 53s - loss: 0.2179 - sparse_categorical_accuracy: 0.9305 - val_loss: 0.1626 - val_sparse_c
ategorical_accuracy: 0.9486 - 53s/epoch - 112ms/step
Epoch 30/100
469/469 - 53s - loss: 0.2174 - sparse_categorical_accuracy: 0.9308 - val_loss: 0.1589 - val_sparse_c
ategorical_accuracy: 0.9505 - 53s/epoch - 112ms/step
Epoch 31/100
469/469 - 53s - loss: 0.2154 - sparse_categorical_accuracy: 0.9323 - val_loss: 0.1548 - val_sparse_c
ategorical_accuracy: 0.9519 - 53s/epoch - 112ms/step
Epoch 32/100
469/469 - 53s - loss: 0.2097 - sparse_categorical_accuracy: 0.9336 - val_loss: 0.1531 - val_sparse_c
ategorical_accuracy: 0.9534 - 53s/epoch - 113ms/step
Epoch 33/100
469/469 - 53s - loss: 0.2108 - sparse_categorical_accuracy: 0.9327 - val_loss: 0.1516 - val_sparse_c
ategorical_accuracy: 0.9531 - 53s/epoch - 113ms/step
Epoch 34/100
469/469 - 53s - loss: 0.2083 - sparse_categorical_accuracy: 0.9336 - val_loss: 0.1518 - val_sparse_c
ategorical_accuracy: 0.9545 - 53s/epoch - 112ms/step
Epoch 35/100
469/469 - 53s - loss: 0.2055 - sparse_categorical_accuracy: 0.9346 - val_loss: 0.1502 - val_sparse_c
ategorical_accuracy: 0.9538 - 53s/epoch - 112ms/step
Epoch 36/100
469/469 - 53s - loss: 0.2024 - sparse_categorical_accuracy: 0.9343 - val_loss: 0.1466 - val_sparse_c
ategorical_accuracy: 0.9561 - 53s/epoch - 113ms/step
Epoch 37/100
469/469 - 53s - loss: 0.2037 - sparse_categorical_accuracy: 0.9346 - val_loss: 0.1496 - val_sparse_c
ategorical_accuracy: 0.9536 - 53s/epoch - 112ms/step
Epoch 38/100
469/469 - 53s - loss: 0.2008 - sparse_categorical_accuracy: 0.9356 - val_loss: 0.1483 - val_sparse_c
ategorical_accuracy: 0.9540 - 53s/epoch - 113ms/step
Epoch 39/100
469/469 - 53s - loss: 0.1960 - sparse_categorical_accuracy: 0.9367 - val_loss: 0.1429 - val_sparse_c
ategorical_accuracy: 0.9573 - 53s/epoch - 112ms/step
Epoch 40/100
469/469 - 53s - loss: 0.1962 - sparse_categorical_accuracy: 0.9373 - val_loss: 0.1417 - val_sparse_c
ategorical_accuracy: 0.9571 - 53s/epoch - 112ms/step
Epoch 41/100
469/469 - 53s - loss: 0.1930 - sparse_categorical_accuracy: 0.9378 - val_loss: 0.1400 - val_sparse_c
ategorical_accuracy: 0.9593 - 53s/epoch - 112ms/step
Epoch 42/100
469/469 - 53s - loss: 0.1929 - sparse_categorical_accuracy: 0.9380 - val_loss: 0.1372 - val_sparse_c
ategorical_accuracy: 0.9582 - 53s/epoch - 112ms/step
Epoch 43/100
469/469 - 53s - loss: 0.1920 - sparse_categorical_accuracy: 0.9388 - val_loss: 0.1398 - val_sparse_c
ategorical_accuracy: 0.9571 - 53s/epoch - 113ms/step
```

```
Epoch 44/100
469/469 - 53s - loss: 0.1900 - sparse_categorical_accuracy: 0.9393 - val_loss: 0.1375 - val_sparse_c
ategorical_accuracy: 0.9567 - 53s/epoch - 113ms/step
Epoch 45/100
469/469 - 53s - loss: 0.1878 - sparse_categorical_accuracy: 0.9398 - val_loss: 0.1401 - val_sparse_c
ategorical_accuracy: 0.9574 - 53s/epoch - 112ms/step
Epoch 46/100
469/469 - 53s - loss: 0.1864 - sparse_categorical_accuracy: 0.9401 - val_loss: 0.1352 - val_sparse_c
ategorical_accuracy: 0.9594 - 53s/epoch - 112ms/step
Epoch 47/100
469/469 - 53s - loss: 0.1877 - sparse_categorical_accuracy: 0.9397 - val_loss: 0.1351 - val_sparse_c
ategorical_accuracy: 0.9596 - 53s/epoch - 112ms/step
Epoch 48/100
469/469 - 53s - loss: 0.1843 - sparse_categorical_accuracy: 0.9413 - val_loss: 0.1351 - val_sparse_c
ategorical_accuracy: 0.9590 - 53s/epoch - 112ms/step
Epoch 49/100
469/469 - 53s - loss: 0.1841 - sparse_categorical_accuracy: 0.9406 - val_loss: 0.1326 - val_sparse_c
ategorical_accuracy: 0.9591 - 53s/epoch - 113ms/step
Epoch 50/100
469/469 - 53s - loss: 0.1820 - sparse_categorical_accuracy: 0.9425 - val_loss: 0.1303 - val_sparse_c
ategorical_accuracy: 0.9610 - 53s/epoch - 113ms/step
Epoch 51/100
469/469 - 53s - loss: 0.1813 - sparse_categorical_accuracy: 0.9416 - val_loss: 0.1308 - val_sparse_c
ategorical_accuracy: 0.9595 - 53s/epoch - 113ms/step
Epoch 52/100
469/469 - 53s - loss: 0.1807 - sparse_categorical_accuracy: 0.9416 - val_loss: 0.1287 - val_sparse_c
ategorical_accuracy: 0.9607 - 53s/epoch - 112ms/step
Epoch 53/100
469/469 - 53s - loss: 0.1792 - sparse_categorical_accuracy: 0.9430 - val_loss: 0.1286 - val_sparse_c
ategorical_accuracy: 0.9605 - 53s/epoch - 112ms/step
Epoch 54/100
469/469 - 53s - loss: 0.1781 - sparse_categorical_accuracy: 0.9427 - val_loss: 0.1311 - val_sparse_c
ategorical_accuracy: 0.9595 - 53s/epoch - 112ms/step
Epoch 55/100
469/469 - 53s - loss: 0.1769 - sparse_categorical_accuracy: 0.9429 - val_loss: 0.1275 - val_sparse_c
ategorical_accuracy: 0.9619 - 53s/epoch - 113ms/step
Epoch 56/100
469/469 - 53s - loss: 0.1792 - sparse_categorical_accuracy: 0.9414 - val_loss: 0.1266 - val_sparse_c
ategorical_accuracy: 0.9626 - 53s/epoch - 113ms/step
Epoch 57/100
469/469 - 53s - loss: 0.1757 - sparse_categorical_accuracy: 0.9440 - val_loss: 0.1281 - val_sparse_c
ategorical_accuracy: 0.9616 - 53s/epoch - 112ms/step
Epoch 58/100
469/469 - 53s - loss: 0.1734 - sparse_categorical_accuracy: 0.9444 - val_loss: 0.1265 - val_sparse_c
ategorical_accuracy: 0.9615 - 53s/epoch - 113ms/step
Epoch 59/100
469/469 - 53s - loss: 0.1738 - sparse_categorical_accuracy: 0.9446 - val_loss: 0.1258 - val_sparse_c
ategorical_accuracy: 0.9628 - 53s/epoch - 113ms/step
Epoch 60/100
469/469 - 53s - loss: 0.1729 - sparse_categorical_accuracy: 0.9442 - val_loss: 0.1245 - val_sparse_c
ategorical_accuracy: 0.9618 - 53s/epoch - 113ms/step
Epoch 61/100
469/469 - 53s - loss: 0.1744 - sparse_categorical_accuracy: 0.9441 - val_loss: 0.1267 - val_sparse_c
ategorical_accuracy: 0.9623 - 53s/epoch - 114ms/step
Epoch 62/100
469/469 - 53s - loss: 0.1712 - sparse_categorical_accuracy: 0.9445 - val_loss: 0.1234 - val_sparse_c
ategorical_accuracy: 0.9623 - 53s/epoch - 113ms/step
Epoch 63/100
469/469 - 53s - loss: 0.1712 - sparse_categorical_accuracy: 0.9450 - val_loss: 0.1209 - val_sparse_c
ategorical_accuracy: 0.9628 - 53s/epoch - 113ms/step
Epoch 64/100
469/469 - 53s - loss: 0.1698 - sparse_categorical_accuracy: 0.9450 - val_loss: 0.1247 - val_sparse_c
ategorical_accuracy: 0.9636 - 53s/epoch - 112ms/step
Epoch 65/100
469/469 - 53s - loss: 0.1698 - sparse_categorical_accuracy: 0.9460 - val_loss: 0.1229 - val_sparse_c
ategorical_accuracy: 0.9615 - 53s/epoch - 113ms/step
Epoch 66/100
469/469 - 53s - loss: 0.1696 - sparse_categorical_accuracy: 0.9451 - val_loss: 0.1225 - val_sparse_c
ategorical_accuracy: 0.9620 - 53s/epoch - 113ms/step
Epoch 67/100
469/469 - 53s - loss: 0.1692 - sparse_categorical_accuracy: 0.9456 - val_loss: 0.1207 - val_sparse_c
ategorical_accuracy: 0.9638 - 53s/epoch - 113ms/step
Epoch 68/100
469/469 - 53s - loss: 0.1676 - sparse_categorical_accuracy: 0.9464 - val_loss: 0.1220 - val_sparse_c
ategorical_accuracy: 0.9637 - 53s/epoch - 112ms/step
Epoch 69/100
469/469 - 53s - loss: 0.1686 - sparse_categorical_accuracy: 0.9458 - val_loss: 0.1210 - val_sparse_c
ategorical_accuracy: 0.9640 - 53s/epoch - 113ms/step
Epoch 70/100
469/469 - 53s - loss: 0.1677 - sparse_categorical_accuracy: 0.9469 - val_loss: 0.1199 - val_sparse_c
ategorical_accuracy: 0.9638 - 53s/epoch - 112ms/step
Epoch 71/100
469/469 - 53s - loss: 0.1668 - sparse_categorical_accuracy: 0.9458 - val_loss: 0.1205 - val_sparse_c
```

```
ategorical_accuracy: 0.9636 - 53s/epoch - 112ms/step
Epoch 72/100
469/469 - 53s - loss: 0.1644 - sparse_categorical_accuracy: 0.9456 - val_loss: 0.1183 - val_sparse_c
ategorical_accuracy: 0.9647 - 53s/epoch - 113ms/step
Epoch 73/100
469/469 - 53s - loss: 0.1631 - sparse_categorical_accuracy: 0.9479 - val_loss: 0.1203 - val_sparse_c
ategorical_accuracy: 0.9631 - 53s/epoch - 113ms/step
Epoch 74/100
469/469 - 53s - loss: 0.1663 - sparse_categorical_accuracy: 0.9468 - val_loss: 0.1167 - val_sparse_c
ategorical_accuracy: 0.9641 - 53s/epoch - 113ms/step
Epoch 75/100
469/469 - 53s - loss: 0.1664 - sparse_categorical_accuracy: 0.9461 - val_loss: 0.1182 - val_sparse_c
ategorical_accuracy: 0.9637 - 53s/epoch - 113ms/step
Epoch 76/100
469/469 - 53s - loss: 0.1656 - sparse_categorical_accuracy: 0.9464 - val_loss: 0.1181 - val_sparse_c
ategorical_accuracy: 0.9650 - 53s/epoch - 113ms/step
Epoch 77/100
469/469 - 53s - loss: 0.1648 - sparse_categorical_accuracy: 0.9467 - val_loss: 0.1183 - val_sparse_c
ategorical_accuracy: 0.9644 - 53s/epoch - 113ms/step
Epoch 78/100
469/469 - 53s - loss: 0.1640 - sparse_categorical_accuracy: 0.9478 - val_loss: 0.1169 - val_sparse_c
ategorical_accuracy: 0.9637 - 53s/epoch - 113ms/step
Epoch 79/100
469/469 - 53s - loss: 0.1624 - sparse_categorical_accuracy: 0.9474 - val_loss: 0.1158 - val_sparse_c
ategorical_accuracy: 0.9658 - 53s/epoch - 113ms/step
Epoch 80/100
469/469 - 53s - loss: 0.1609 - sparse_categorical_accuracy: 0.9488 - val_loss: 0.1158 - val_sparse_c
ategorical_accuracy: 0.9662 - 53s/epoch - 113ms/step
Epoch 81/100
469/469 - 53s - loss: 0.1626 - sparse_categorical_accuracy: 0.9478 - val_loss: 0.1159 - val_sparse_c
ategorical_accuracy: 0.9652 - 53s/epoch - 113ms/step
Epoch 82/100
469/469 - 53s - loss: 0.1620 - sparse_categorical_accuracy: 0.9480 - val_loss: 0.1166 - val_sparse_c
ategorical_accuracy: 0.9645 - 53s/epoch - 113ms/step
Epoch 83/100
469/469 - 53s - loss: 0.1632 - sparse_categorical_accuracy: 0.9482 - val_loss: 0.1150 - val_sparse_c
ategorical_accuracy: 0.9650 - 53s/epoch - 113ms/step
Epoch 84/100
469/469 - 57s - loss: 0.1595 - sparse_categorical_accuracy: 0.9488 - val_loss: 0.1138 - val_sparse_c
ategorical_accuracy: 0.9655 - 57s/epoch - 122ms/step
Epoch 85/100
469/469 - 53s - loss: 0.1609 - sparse_categorical_accuracy: 0.9493 - val_loss: 0.1140 - val_sparse_c
ategorical_accuracy: 0.9648 - 53s/epoch - 114ms/step
Epoch 86/100
469/469 - 53s - loss: 0.1610 - sparse_categorical_accuracy: 0.9486 - val_loss: 0.1174 - val_sparse_c
ategorical_accuracy: 0.9644 - 53s/epoch - 113ms/step
Epoch 87/100
469/469 - 53s - loss: 0.1592 - sparse_categorical_accuracy: 0.9490 - val_loss: 0.1160 - val_sparse_c
ategorical_accuracy: 0.9656 - 53s/epoch - 113ms/step
Epoch 88/100
469/469 - 53s - loss: 0.1617 - sparse_categorical_accuracy: 0.9484 - val_loss: 0.1133 - val_sparse_c
ategorical_accuracy: 0.9663 - 53s/epoch - 112ms/step
Epoch 89/100
469/469 - 53s - loss: 0.1590 - sparse_categorical_accuracy: 0.9498 - val_loss: 0.1162 - val_sparse_c
ategorical_accuracy: 0.9653 - 53s/epoch - 113ms/step
Epoch 90/100
469/469 - 53s - loss: 0.1567 - sparse_categorical_accuracy: 0.9496 - val_loss: 0.1146 - val_sparse_c
ategorical_accuracy: 0.9658 - 53s/epoch - 113ms/step
Epoch 91/100
469/469 - 53s - loss: 0.1565 - sparse_categorical_accuracy: 0.9498 - val_loss: 0.1148 - val_sparse_c
ategorical_accuracy: 0.9664 - 53s/epoch - 112ms/step
Epoch 92/100
469/469 - 53s - loss: 0.1559 - sparse_categorical_accuracy: 0.9492 - val_loss: 0.1140 - val_sparse_c
ategorical_accuracy: 0.9654 - 53s/epoch - 112ms/step
Epoch 93/100
469/469 - 53s - loss: 0.1577 - sparse_categorical_accuracy: 0.9496 - val_loss: 0.1135 - val_sparse_c
ategorical_accuracy: 0.9664 - 53s/epoch - 112ms/step
Epoch 94/100
469/469 - 53s - loss: 0.1562 - sparse_categorical_accuracy: 0.9496 - val_loss: 0.1122 - val_sparse_c
ategorical_accuracy: 0.9663 - 53s/epoch - 113ms/step
Epoch 95/100
469/469 - 53s - loss: 0.1556 - sparse_categorical_accuracy: 0.9496 - val_loss: 0.1135 - val_sparse_c
ategorical_accuracy: 0.9654 - 53s/epoch - 113ms/step
Epoch 96/100
469/469 - 53s - loss: 0.1554 - sparse_categorical_accuracy: 0.9505 - val_loss: 0.1136 - val_sparse_c
ategorical_accuracy: 0.9659 - 53s/epoch - 112ms/step
Epoch 97/100
469/469 - 53s - loss: 0.1521 - sparse_categorical_accuracy: 0.9517 - val_loss: 0.1123 - val_sparse_c
ategorical_accuracy: 0.9676 - 53s/epoch - 112ms/step
Epoch 98/100
469/469 - 53s - loss: 0.1557 - sparse_categorical_accuracy: 0.9499 - val_loss: 0.1130 - val_sparse_c
ategorical_accuracy: 0.9665 - 53s/epoch - 112ms/step
Epoch 99/100
```

```
469/469 - 53s - loss: 0.1556 - sparse_categorical_accuracy: 0.9494 - val_loss: 0.1099 - val_sparse_c
ategorical_accuracy: 0.9667 - 53s/epoch - 112ms/step
Epoch 100/100
469/469 - 53s - loss: 0.1571 - sparse_categorical_accuracy: 0.9498 - val_loss: 0.1114 - val_sparse_c
ategorical_accuracy: 0.9663 - 53s/epoch - 112ms/step
```

In [24]:

```
model4.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 4s - loss: 0.1114 - sparse_categorical_accuracy: 0.9663 - 4s/epoch - 13ms/step
```

Out[24]:

```
[0.11144724488258362, 0.9663000106811523]
```

Model5: `adam` optimizer with learning rate=e^{-3}, `random_uniform` initializer, `dropout` regularization with rate=0.2.

In [25]:

```
input_layer = Input(shape=(28, 28,1))
x = inception_module(input_layer,
                     filters_1x1=1,
                     filters_3x3_reduce=1,
                     filters_3x3=1,
                     filters_5x5_reduce=1,
                     filters_5x5=1,
                     filters_pool_proj=1,
                     name='inception_3a')
x = tf.keras.layers.BatchNormalization()(x)
x = MaxPool2D((3, 3))(x)
x = tf.keras.layers.Flatten()(x)
x = Dropout(0.2)(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform')(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform')(x)
x = tf.keras.layers.BatchNormalization()(x)
x = Dense(10, activation='softmax', kernel_initializer='random_uniform')(x)
model5 = Model(input_layer, x, name='inception_v1')
```

In [26]:

```
model5.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [27]:

```
history5 = model5.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 54s - loss: 0.8432 - sparse_categorical_accuracy: 0.8047 - val_loss: 0.2481 - val_sparse_c
ategorical_accuracy: 0.9427 - 54s/epoch - 115ms/step
Epoch 2/100
469/469 - 52s - loss: 0.2471 - sparse_categorical_accuracy: 0.9296 - val_loss: 0.1435 - val_sparse_c
ategorical_accuracy: 0.9582 - 52s/epoch - 112ms/step
Epoch 3/100
469/469 - 53s - loss: 0.1998 - sparse_categorical_accuracy: 0.9379 - val_loss: 0.1294 - val_sparse_c
ategorical_accuracy: 0.9611 - 53s/epoch - 112ms/step
Epoch 4/100
469/469 - 53s - loss: 0.1800 - sparse_categorical_accuracy: 0.9438 - val_loss: 0.1117 - val_sparse_c
ategorical_accuracy: 0.9658 - 53s/epoch - 112ms/step
Epoch 5/100
469/469 - 53s - loss: 0.1727 - sparse_categorical_accuracy: 0.9458 - val_loss: 0.1038 - val_sparse_c
ategorical_accuracy: 0.9680 - 53s/epoch - 112ms/step
Epoch 6/100
469/469 - 53s - loss: 0.1626 - sparse_categorical_accuracy: 0.9492 - val_loss: 0.1076 - val_sparse_c
ategorical_accuracy: 0.9649 - 53s/epoch - 112ms/step
Epoch 7/100
469/469 - 56s - loss: 0.1601 - sparse_categorical_accuracy: 0.9491 - val_loss: 0.1027 - val_sparse_c
ategorical_accuracy: 0.9667 - 56s/epoch - 120ms/step
Epoch 8/100
469/469 - 53s - loss: 0.1503 - sparse_categorical_accuracy: 0.9522 - val_loss: 0.1003 - val_sparse_c
ategorical_accuracy: 0.9672 - 53s/epoch - 112ms/step
Epoch 9/100
469/469 - 54s - loss: 0.1465 - sparse_categorical_accuracy: 0.9539 - val_loss: 0.0923 - val_sparse_c
ategorical_accuracy: 0.9706 - 54s/epoch - 114ms/step
```

```
Epoch 10/100
469/469 - 53s - loss: 0.1434 - sparse_categorical_accuracy: 0.9538 - val_loss: 0.0875 - val_sparse_c
ategorical_accuracy: 0.9718 - 53s/epoch - 112ms/step
Epoch 11/100
469/469 - 53s - loss: 0.1397 - sparse_categorical_accuracy: 0.9548 - val_loss: 0.0902 - val_sparse_c
ategorical_accuracy: 0.9705 - 53s/epoch - 113ms/step
Epoch 12/100
469/469 - 53s - loss: 0.1361 - sparse_categorical_accuracy: 0.9571 - val_loss: 0.0890 - val_sparse_c
ategorical_accuracy: 0.9708 - 53s/epoch - 112ms/step
Epoch 13/100
469/469 - 53s - loss: 0.1337 - sparse_categorical_accuracy: 0.9575 - val_loss: 0.0819 - val_sparse_c
ategorical_accuracy: 0.9733 - 53s/epoch - 112ms/step
Epoch 14/100
469/469 - 53s - loss: 0.1310 - sparse_categorical_accuracy: 0.9580 - val_loss: 0.0893 - val_sparse_c
ategorical_accuracy: 0.9712 - 53s/epoch - 112ms/step
Epoch 15/100
469/469 - 53s - loss: 0.1276 - sparse_categorical_accuracy: 0.9586 - val_loss: 0.0864 - val_sparse_c
ategorical_accuracy: 0.9711 - 53s/epoch - 112ms/step
Epoch 16/100
469/469 - 53s - loss: 0.1291 - sparse_categorical_accuracy: 0.9584 - val_loss: 0.0954 - val_sparse_c
ategorical_accuracy: 0.9691 - 53s/epoch - 112ms/step
Epoch 17/100
469/469 - 53s - loss: 0.1269 - sparse_categorical_accuracy: 0.9587 - val_loss: 0.0829 - val_sparse_c
ategorical_accuracy: 0.9730 - 53s/epoch - 113ms/step
Epoch 18/100
469/469 - 53s - loss: 0.1221 - sparse_categorical_accuracy: 0.9603 - val_loss: 0.0758 - val_sparse_c
ategorical_accuracy: 0.9752 - 53s/epoch - 112ms/step
Epoch 19/100
469/469 - 53s - loss: 0.1239 - sparse_categorical_accuracy: 0.9596 - val_loss: 0.0798 - val_sparse_c
ategorical_accuracy: 0.9750 - 53s/epoch - 113ms/step
Epoch 20/100
469/469 - 53s - loss: 0.1234 - sparse_categorical_accuracy: 0.9601 - val_loss: 0.0835 - val_sparse_c
ategorical_accuracy: 0.9743 - 53s/epoch - 112ms/step
Epoch 21/100
469/469 - 53s - loss: 0.1208 - sparse_categorical_accuracy: 0.9610 - val_loss: 0.0765 - val_sparse_c
ategorical_accuracy: 0.9760 - 53s/epoch - 113ms/step
Epoch 22/100
469/469 - 53s - loss: 0.1206 - sparse_categorical_accuracy: 0.9613 - val_loss: 0.0772 - val_sparse_c
ategorical_accuracy: 0.9753 - 53s/epoch - 112ms/step
Epoch 23/100
469/469 - 53s - loss: 0.1171 - sparse_categorical_accuracy: 0.9621 - val_loss: 0.0789 - val_sparse_c
ategorical_accuracy: 0.9740 - 53s/epoch - 113ms/step
Epoch 24/100
469/469 - 53s - loss: 0.1164 - sparse_categorical_accuracy: 0.9614 - val_loss: 0.0737 - val_sparse_c
ategorical_accuracy: 0.9764 - 53s/epoch - 113ms/step
Epoch 25/100
469/469 - 53s - loss: 0.1160 - sparse_categorical_accuracy: 0.9628 - val_loss: 0.0708 - val_sparse_c
ategorical_accuracy: 0.9785 - 53s/epoch - 113ms/step
Epoch 26/100
469/469 - 53s - loss: 0.1164 - sparse_categorical_accuracy: 0.9622 - val_loss: 0.0730 - val_sparse_c
ategorical_accuracy: 0.9764 - 53s/epoch - 113ms/step
Epoch 27/100
469/469 - 53s - loss: 0.1176 - sparse_categorical_accuracy: 0.9617 - val_loss: 0.0715 - val_sparse_c
ategorical_accuracy: 0.9775 - 53s/epoch - 112ms/step
Epoch 28/100
469/469 - 53s - loss: 0.1140 - sparse_categorical_accuracy: 0.9635 - val_loss: 0.0771 - val_sparse_c
ategorical_accuracy: 0.9749 - 53s/epoch - 113ms/step
Epoch 29/100
469/469 - 53s - loss: 0.1110 - sparse_categorical_accuracy: 0.9638 - val_loss: 0.0765 - val_sparse_c
ategorical_accuracy: 0.9753 - 53s/epoch - 113ms/step
Epoch 30/100
469/469 - 53s - loss: 0.1129 - sparse_categorical_accuracy: 0.9632 - val_loss: 0.0710 - val_sparse_c
ategorical_accuracy: 0.9770 - 53s/epoch - 112ms/step
Epoch 31/100
469/469 - 53s - loss: 0.1128 - sparse_categorical_accuracy: 0.9633 - val_loss: 0.0700 - val_sparse_c
ategorical_accuracy: 0.9777 - 53s/epoch - 112ms/step
Epoch 32/100
469/469 - 53s - loss: 0.1108 - sparse_categorical_accuracy: 0.9644 - val_loss: 0.0692 - val_sparse_c
ategorical_accuracy: 0.9779 - 53s/epoch - 113ms/step
Epoch 33/100
469/469 - 53s - loss: 0.1120 - sparse_categorical_accuracy: 0.9634 - val_loss: 0.0767 - val_sparse_c
ategorical_accuracy: 0.9750 - 53s/epoch - 112ms/step
Epoch 34/100
469/469 - 53s - loss: 0.1095 - sparse_categorical_accuracy: 0.9642 - val_loss: 0.0765 - val_sparse_c
ategorical_accuracy: 0.9764 - 53s/epoch - 113ms/step
Epoch 35/100
469/469 - 53s - loss: 0.1119 - sparse_categorical_accuracy: 0.9646 - val_loss: 0.0778 - val_sparse_c
ategorical_accuracy: 0.9749 - 53s/epoch - 112ms/step
Epoch 36/100
469/469 - 53s - loss: 0.1110 - sparse_categorical_accuracy: 0.9637 - val_loss: 0.0715 - val_sparse_c
ategorical_accuracy: 0.9757 - 53s/epoch - 112ms/step
Epoch 37/100
469/469 - 53s - loss: 0.1082 - sparse_categorical_accuracy: 0.9644 - val_loss: 0.0749 - val_sparse_c
```

```
ategorical_accuracy: 0.9758 - 53s/epoch - 113ms/step
Epoch 38/100
469/469 - 53s - loss: 0.1090 - sparse_categorical_accuracy: 0.9646 - val_loss: 0.0681 - val_sparse_c
ategorical_accuracy: 0.9772 - 53s/epoch - 112ms/step
Epoch 39/100
469/469 - 53s - loss: 0.1089 - sparse_categorical_accuracy: 0.9653 - val_loss: 0.0695 - val_sparse_c
ategorical_accuracy: 0.9771 - 53s/epoch - 112ms/step
Epoch 40/100
469/469 - 53s - loss: 0.1054 - sparse_categorical_accuracy: 0.9656 - val_loss: 0.0665 - val_sparse_c
ategorical_accuracy: 0.9781 - 53s/epoch - 112ms/step
Epoch 41/100
469/469 - 53s - loss: 0.1091 - sparse_categorical_accuracy: 0.9652 - val_loss: 0.0671 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 113ms/step
Epoch 42/100
469/469 - 53s - loss: 0.1077 - sparse_categorical_accuracy: 0.9651 - val_loss: 0.0676 - val_sparse_c
ategorical_accuracy: 0.9784 - 53s/epoch - 112ms/step
Epoch 43/100
469/469 - 53s - loss: 0.1059 - sparse_categorical_accuracy: 0.9653 - val_loss: 0.0653 - val_sparse_c
ategorical_accuracy: 0.9780 - 53s/epoch - 112ms/step
Epoch 44/100
469/469 - 53s - loss: 0.1068 - sparse_categorical_accuracy: 0.9655 - val_loss: 0.0710 - val_sparse_c
ategorical_accuracy: 0.9771 - 53s/epoch - 112ms/step
Epoch 45/100
469/469 - 53s - loss: 0.1065 - sparse_categorical_accuracy: 0.9650 - val_loss: 0.0685 - val_sparse_c
ategorical_accuracy: 0.9784 - 53s/epoch - 113ms/step
Epoch 46/100
469/469 - 53s - loss: 0.1078 - sparse_categorical_accuracy: 0.9657 - val_loss: 0.0679 - val_sparse_c
ategorical_accuracy: 0.9781 - 53s/epoch - 113ms/step
Epoch 47/100
469/469 - 52s - loss: 0.1053 - sparse_categorical_accuracy: 0.9657 - val_loss: 0.0693 - val_sparse_c
ategorical_accuracy: 0.9766 - 52s/epoch - 112ms/step
Epoch 48/100
469/469 - 53s - loss: 0.1053 - sparse_categorical_accuracy: 0.9653 - val_loss: 0.0678 - val_sparse_c
ategorical_accuracy: 0.9785 - 53s/epoch - 113ms/step
Epoch 49/100
469/469 - 53s - loss: 0.1046 - sparse_categorical_accuracy: 0.9660 - val_loss: 0.0677 - val_sparse_c
ategorical_accuracy: 0.9781 - 53s/epoch - 113ms/step
Epoch 50/100
469/469 - 53s - loss: 0.1037 - sparse_categorical_accuracy: 0.9661 - val_loss: 0.0728 - val_sparse_c
ategorical_accuracy: 0.9767 - 53s/epoch - 113ms/step
Epoch 51/100
469/469 - 53s - loss: 0.1039 - sparse_categorical_accuracy: 0.9665 - val_loss: 0.0689 - val_sparse_c
ategorical_accuracy: 0.9774 - 53s/epoch - 113ms/step
Epoch 52/100
469/469 - 53s - loss: 0.1036 - sparse_categorical_accuracy: 0.9667 - val_loss: 0.0681 - val_sparse_c
ategorical_accuracy: 0.9779 - 53s/epoch - 112ms/step
Epoch 53/100
469/469 - 53s - loss: 0.1025 - sparse_categorical_accuracy: 0.9668 - val_loss: 0.0668 - val_sparse_c
ategorical_accuracy: 0.9782 - 53s/epoch - 113ms/step
Epoch 54/100
469/469 - 53s - loss: 0.1038 - sparse_categorical_accuracy: 0.9663 - val_loss: 0.0676 - val_sparse_c
ategorical_accuracy: 0.9775 - 53s/epoch - 113ms/step
Epoch 55/100
469/469 - 53s - loss: 0.1016 - sparse_categorical_accuracy: 0.9673 - val_loss: 0.0654 - val_sparse_c
ategorical_accuracy: 0.9786 - 53s/epoch - 112ms/step
Epoch 56/100
469/469 - 53s - loss: 0.1043 - sparse_categorical_accuracy: 0.9660 - val_loss: 0.0640 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 112ms/step
Epoch 57/100
469/469 - 53s - loss: 0.1021 - sparse_categorical_accuracy: 0.9664 - val_loss: 0.0632 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 113ms/step
Epoch 58/100
469/469 - 53s - loss: 0.1039 - sparse_categorical_accuracy: 0.9663 - val_loss: 0.0650 - val_sparse_c
ategorical_accuracy: 0.9785 - 53s/epoch - 112ms/step
Epoch 59/100
469/469 - 53s - loss: 0.1020 - sparse_categorical_accuracy: 0.9669 - val_loss: 0.0683 - val_sparse_c
ategorical_accuracy: 0.9776 - 53s/epoch - 112ms/step
Epoch 60/100
469/469 - 53s - loss: 0.1037 - sparse_categorical_accuracy: 0.9659 - val_loss: 0.0662 - val_sparse_c
ategorical_accuracy: 0.9788 - 53s/epoch - 112ms/step
Epoch 61/100
469/469 - 53s - loss: 0.1011 - sparse_categorical_accuracy: 0.9675 - val_loss: 0.0696 - val_sparse_c
ategorical_accuracy: 0.9776 - 53s/epoch - 112ms/step
Epoch 62/100
469/469 - 53s - loss: 0.1025 - sparse_categorical_accuracy: 0.9668 - val_loss: 0.0640 - val_sparse_c
ategorical_accuracy: 0.9790 - 53s/epoch - 113ms/step
Epoch 63/100
469/469 - 53s - loss: 0.1001 - sparse_categorical_accuracy: 0.9675 - val_loss: 0.0631 - val_sparse_c
ategorical_accuracy: 0.9791 - 53s/epoch - 113ms/step
Epoch 64/100
469/469 - 53s - loss: 0.1021 - sparse_categorical_accuracy: 0.9666 - val_loss: 0.0627 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 113ms/step
Epoch 65/100
```

```
469/469 - 53s - loss: 0.0982 - sparse_categorical_accuracy: 0.9678 - val_loss: 0.0642 - val_sparse_c
ategorical_accuracy: 0.9791 - 53s/epoch - 112ms/step
Epoch 66/100
469/469 - 53s - loss: 0.1006 - sparse_categorical_accuracy: 0.9673 - val_loss: 0.0611 - val_sparse_c
ategorical_accuracy: 0.9818 - 53s/epoch - 112ms/step
Epoch 67/100
469/469 - 53s - loss: 0.0998 - sparse_categorical_accuracy: 0.9677 - val_loss: 0.0628 - val_sparse_c
ategorical_accuracy: 0.9795 - 53s/epoch - 112ms/step
Epoch 68/100
469/469 - 53s - loss: 0.0988 - sparse_categorical_accuracy: 0.9686 - val_loss: 0.0614 - val_sparse_c
ategorical_accuracy: 0.9797 - 53s/epoch - 113ms/step
Epoch 69/100
469/469 - 53s - loss: 0.0984 - sparse_categorical_accuracy: 0.9673 - val_loss: 0.0640 - val_sparse_c
ategorical_accuracy: 0.9783 - 53s/epoch - 113ms/step
Epoch 70/100
469/469 - 53s - loss: 0.0979 - sparse_categorical_accuracy: 0.9678 - val_loss: 0.0617 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 113ms/step
Epoch 71/100
469/469 - 53s - loss: 0.0979 - sparse_categorical_accuracy: 0.9681 - val_loss: 0.0655 - val_sparse_c
ategorical_accuracy: 0.9784 - 53s/epoch - 112ms/step
Epoch 72/100
469/469 - 53s - loss: 0.0991 - sparse_categorical_accuracy: 0.9670 - val_loss: 0.0662 - val_sparse_c
ategorical_accuracy: 0.9783 - 53s/epoch - 112ms/step
Epoch 73/100
469/469 - 53s - loss: 0.0962 - sparse_categorical_accuracy: 0.9689 - val_loss: 0.0666 - val_sparse_c
ategorical_accuracy: 0.9785 - 53s/epoch - 113ms/step
Epoch 74/100
469/469 - 53s - loss: 0.0989 - sparse_categorical_accuracy: 0.9673 - val_loss: 0.0662 - val_sparse_c
ategorical_accuracy: 0.9780 - 53s/epoch - 113ms/step
Epoch 75/100
469/469 - 53s - loss: 0.0986 - sparse_categorical_accuracy: 0.9683 - val_loss: 0.0696 - val_sparse_c
ategorical_accuracy: 0.9762 - 53s/epoch - 112ms/step
Epoch 76/100
469/469 - 53s - loss: 0.0977 - sparse_categorical_accuracy: 0.9681 - val_loss: 0.0642 - val_sparse_c
ategorical_accuracy: 0.9792 - 53s/epoch - 112ms/step
Epoch 77/100
469/469 - 53s - loss: 0.0961 - sparse_categorical_accuracy: 0.9688 - val_loss: 0.0626 - val_sparse_c
ategorical_accuracy: 0.9805 - 53s/epoch - 112ms/step
Epoch 78/100
469/469 - 53s - loss: 0.0968 - sparse_categorical_accuracy: 0.9685 - val_loss: 0.0623 - val_sparse_c
ategorical_accuracy: 0.9799 - 53s/epoch - 112ms/step
Epoch 79/100
469/469 - 53s - loss: 0.0958 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0603 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 112ms/step
Epoch 80/100
469/469 - 53s - loss: 0.0952 - sparse_categorical_accuracy: 0.9685 - val_loss: 0.0623 - val_sparse_c
ategorical_accuracy: 0.9796 - 53s/epoch - 113ms/step
Epoch 81/100
469/469 - 53s - loss: 0.0966 - sparse_categorical_accuracy: 0.9688 - val_loss: 0.0604 - val_sparse_c
ategorical_accuracy: 0.9810 - 53s/epoch - 113ms/step
Epoch 82/100
469/469 - 53s - loss: 0.0969 - sparse_categorical_accuracy: 0.9681 - val_loss: 0.0658 - val_sparse_c
ategorical_accuracy: 0.9784 - 53s/epoch - 112ms/step
Epoch 83/100
469/469 - 53s - loss: 0.0969 - sparse_categorical_accuracy: 0.9686 - val_loss: 0.0630 - val_sparse_c
ategorical_accuracy: 0.9789 - 53s/epoch - 112ms/step
Epoch 84/100
469/469 - 52s - loss: 0.0970 - sparse_categorical_accuracy: 0.9678 - val_loss: 0.0623 - val_sparse_c
ategorical_accuracy: 0.9798 - 52s/epoch - 112ms/step
Epoch 85/100
469/469 - 53s - loss: 0.0969 - sparse_categorical_accuracy: 0.9686 - val_loss: 0.0665 - val_sparse_c
ategorical_accuracy: 0.9783 - 53s/epoch - 112ms/step
Epoch 86/100
469/469 - 53s - loss: 0.0964 - sparse_categorical_accuracy: 0.9680 - val_loss: 0.0641 - val_sparse_c
ategorical_accuracy: 0.9792 - 53s/epoch - 112ms/step
Epoch 87/100
469/469 - 53s - loss: 0.0958 - sparse_categorical_accuracy: 0.9682 - val_loss: 0.0628 - val_sparse_c
ategorical_accuracy: 0.9799 - 53s/epoch - 112ms/step
Epoch 88/100
469/469 - 53s - loss: 0.0981 - sparse_categorical_accuracy: 0.9678 - val_loss: 0.0624 - val_sparse_c
ategorical_accuracy: 0.9802 - 53s/epoch - 112ms/step
Epoch 89/100
469/469 - 53s - loss: 0.0966 - sparse_categorical_accuracy: 0.9681 - val_loss: 0.0610 - val_sparse_c
ategorical_accuracy: 0.9808 - 53s/epoch - 112ms/step
Epoch 90/100
469/469 - 53s - loss: 0.0948 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0611 - val_sparse_c
ategorical_accuracy: 0.9795 - 53s/epoch - 113ms/step
Epoch 91/100
469/469 - 53s - loss: 0.0964 - sparse_categorical_accuracy: 0.9689 - val_loss: 0.0608 - val_sparse_c
ategorical_accuracy: 0.9807 - 53s/epoch - 112ms/step
Epoch 92/100
469/469 - 53s - loss: 0.0933 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0623 - val_sparse_c
ategorical_accuracy: 0.9791 - 53s/epoch - 113ms/step
```

```
Epoch 93/100
469/469 - 53s - loss: 0.0969 - sparse_categorical_accuracy: 0.9679 - val_loss: 0.0637 - val_sparse_c
ategorical_accuracy: 0.9788 - 53s/epoch - 112ms/step
Epoch 94/100
469/469 - 53s - loss: 0.0968 - sparse_categorical_accuracy: 0.9688 - val_loss: 0.0647 - val_sparse_c
ategorical_accuracy: 0.9787 - 53s/epoch - 112ms/step
Epoch 95/100
469/469 - 52s - loss: 0.0937 - sparse_categorical_accuracy: 0.9690 - val_loss: 0.0629 - val_sparse_c
ategorical_accuracy: 0.9794 - 52s/epoch - 112ms/step
Epoch 96/100
469/469 - 53s - loss: 0.0950 - sparse_categorical_accuracy: 0.9692 - val_loss: 0.0627 - val_sparse_c
ategorical_accuracy: 0.9793 - 53s/epoch - 112ms/step
Epoch 97/100
469/469 - 53s - loss: 0.0945 - sparse_categorical_accuracy: 0.9690 - val_loss: 0.0619 - val_sparse_c
ategorical_accuracy: 0.9808 - 53s/epoch - 112ms/step
Epoch 98/100
469/469 - 53s - loss: 0.0921 - sparse_categorical_accuracy: 0.9699 - val_loss: 0.0635 - val_sparse_c
ategorical_accuracy: 0.9788 - 53s/epoch - 112ms/step
Epoch 99/100
469/469 - 53s - loss: 0.0972 - sparse_categorical_accuracy: 0.9688 - val_loss: 0.0603 - val_sparse_c
ategorical_accuracy: 0.9804 - 53s/epoch - 113ms/step
Epoch 100/100
469/469 - 53s - loss: 0.0958 - sparse_categorical_accuracy: 0.9687 - val_loss: 0.0614 - val_sparse_c
ategorical_accuracy: 0.9794 - 53s/epoch - 113ms/step
```

In [28]:

```
model5.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 4s - loss: 0.0614 - sparse_categorical_accuracy: 0.9794 - 4s/epoch - 13ms/step
```

Out[28]:

```
[0.0613691546022892, 0.9793999791145325]
```

Model6: `adam` optimizer with learning rate=e^{-4}, `random_uniform` initializer, `dropout` regularization with rate=0.2.

In [29]:

```
input_layer = Input(shape=(28, 28,1))
x = inception_module(input_layer,
                     filters_1x1=1,
                     filters_3x3_reduce=1,
                     filters_3x3=1,
                     filters_5x5_reduce=1,
                     filters_5x5=1,
                     filters_pool_proj=1,
                     name='inception_3a')
x = tf.keras.layers.BatchNormalization()(x)
x = MaxPool2D((3, 3))(x)
x = tf.keras.layers.Flatten()(x)
x = Dropout(0.2)(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform')(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform')(x)
x = tf.keras.layers.BatchNormalization()(x)
x = Dense(10, activation='softmax', kernel_initializer='random_uniform')(x)
model6 = Model(input_layer, x, name='inception_v1')
```

In [30]:

```
model6.compile(optimizer=tf.keras.optimizers.Adam(0.0001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [31]:

```
history6 = model6.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 54s - loss: 2.0855 - sparse_categorical_accuracy: 0.4488 - val_loss: 1.8832 - val_sparse_c
ategorical_accuracy: 0.6610 - 54s/epoch - 115ms/step
Epoch 2/100
469/469 - 53s - loss: 1.6718 - sparse_categorical_accuracy: 0.6506 - val_loss: 1.4267 - val_sparse_c
ategorical_accuracy: 0.7515 - 53s/epoch - 113ms/step
Epoch 3/100
469/469 - 53s - loss: 1.3215 - sparse_categorical_accuracy: 0.7239 - val_loss: 1.0785 - val_sparse_c
```

```
ategorical_accuracy: 0.8166 - 53s/epoch - 112ms/step
Epoch 4/100
469/469 - 53s - loss: 1.0307 - sparse_categorical_accuracy: 0.7755 - val_loss: 0.8302 - val_sparse_c
ategorical_accuracy: 0.8522 - 53s/epoch - 113ms/step
Epoch 5/100
469/469 - 53s - loss: 0.8388 - sparse_categorical_accuracy: 0.8053 - val_loss: 0.6562 - val_sparse_c
ategorical_accuracy: 0.8659 - 53s/epoch - 112ms/step
Epoch 6/100
469/469 - 53s - loss: 0.6993 - sparse_categorical_accuracy: 0.8195 - val_loss: 0.5278 - val_sparse_c
ategorical_accuracy: 0.8772 - 53s/epoch - 112ms/step
Epoch 7/100
469/469 - 53s - loss: 0.6142 - sparse_categorical_accuracy: 0.8294 - val_loss: 0.4651 - val_sparse_c
ategorical_accuracy: 0.8841 - 53s/epoch - 113ms/step
Epoch 8/100
469/469 - 53s - loss: 0.5653 - sparse_categorical_accuracy: 0.8353 - val_loss: 0.4191 - val_sparse_c
ategorical_accuracy: 0.8912 - 53s/epoch - 114ms/step
Epoch 9/100
469/469 - 53s - loss: 0.5342 - sparse_categorical_accuracy: 0.8395 - val_loss: 0.3901 - val_sparse_c
ategorical_accuracy: 0.8932 - 53s/epoch - 113ms/step
Epoch 10/100
469/469 - 53s - loss: 0.5130 - sparse_categorical_accuracy: 0.8440 - val_loss: 0.3730 - val_sparse_c
ategorical_accuracy: 0.8950 - 53s/epoch - 113ms/step
Epoch 11/100
469/469 - 53s - loss: 0.4925 - sparse_categorical_accuracy: 0.8468 - val_loss: 0.3521 - val_sparse_c
ategorical_accuracy: 0.9011 - 53s/epoch - 113ms/step
Epoch 12/100
469/469 - 53s - loss: 0.4842 - sparse_categorical_accuracy: 0.8464 - val_loss: 0.3451 - val_sparse_c
ategorical_accuracy: 0.9017 - 53s/epoch - 113ms/step
Epoch 13/100
469/469 - 53s - loss: 0.4734 - sparse_categorical_accuracy: 0.8513 - val_loss: 0.3348 - val_sparse_c
ategorical_accuracy: 0.9015 - 53s/epoch - 114ms/step
Epoch 14/100
469/469 - 53s - loss: 0.4612 - sparse_categorical_accuracy: 0.8531 - val_loss: 0.3244 - val_sparse_c
ategorical_accuracy: 0.9048 - 53s/epoch - 114ms/step
Epoch 15/100
469/469 - 53s - loss: 0.4553 - sparse_categorical_accuracy: 0.8544 - val_loss: 0.3211 - val_sparse_c
ategorical_accuracy: 0.9051 - 53s/epoch - 113ms/step
Epoch 16/100
469/469 - 53s - loss: 0.4479 - sparse_categorical_accuracy: 0.8541 - val_loss: 0.3109 - val_sparse_c
ategorical_accuracy: 0.9080 - 53s/epoch - 113ms/step
Epoch 17/100
469/469 - 53s - loss: 0.4383 - sparse_categorical_accuracy: 0.8589 - val_loss: 0.3030 - val_sparse_c
ategorical_accuracy: 0.9077 - 53s/epoch - 113ms/step
Epoch 18/100
469/469 - 53s - loss: 0.4352 - sparse_categorical_accuracy: 0.8596 - val_loss: 0.2985 - val_sparse_c
ategorical_accuracy: 0.9096 - 53s/epoch - 113ms/step
Epoch 19/100
469/469 - 53s - loss: 0.4286 - sparse_categorical_accuracy: 0.8599 - val_loss: 0.2936 - val_sparse_c
ategorical_accuracy: 0.9102 - 53s/epoch - 114ms/step
Epoch 20/100
469/469 - 53s - loss: 0.4285 - sparse_categorical_accuracy: 0.8608 - val_loss: 0.2929 - val_sparse_c
ategorical_accuracy: 0.9095 - 53s/epoch - 113ms/step
Epoch 21/100
469/469 - 53s - loss: 0.4194 - sparse_categorical_accuracy: 0.8641 - val_loss: 0.2874 - val_sparse_c
ategorical_accuracy: 0.9116 - 53s/epoch - 113ms/step
Epoch 22/100
469/469 - 53s - loss: 0.4229 - sparse_categorical_accuracy: 0.8625 - val_loss: 0.2858 - val_sparse_c
ategorical_accuracy: 0.9120 - 53s/epoch - 113ms/step
Epoch 23/100
469/469 - 53s - loss: 0.4164 - sparse_categorical_accuracy: 0.8638 - val_loss: 0.2829 - val_sparse_c
ategorical_accuracy: 0.9127 - 53s/epoch - 113ms/step
Epoch 24/100
469/469 - 53s - loss: 0.4129 - sparse_categorical_accuracy: 0.8663 - val_loss: 0.2919 - val_sparse_c
ategorical_accuracy: 0.9073 - 53s/epoch - 113ms/step
Epoch 25/100
469/469 - 53s - loss: 0.4074 - sparse_categorical_accuracy: 0.8669 - val_loss: 0.2788 - val_sparse_c
ategorical_accuracy: 0.9121 - 53s/epoch - 114ms/step
Epoch 26/100
469/469 - 53s - loss: 0.4066 - sparse_categorical_accuracy: 0.8666 - val_loss: 0.2714 - val_sparse_c
ategorical_accuracy: 0.9151 - 53s/epoch - 113ms/step
Epoch 27/100
469/469 - 53s - loss: 0.4023 - sparse_categorical_accuracy: 0.8695 - val_loss: 0.2664 - val_sparse_c
ategorical_accuracy: 0.9166 - 53s/epoch - 113ms/step
Epoch 28/100
469/469 - 53s - loss: 0.3986 - sparse_categorical_accuracy: 0.8692 - val_loss: 0.2745 - val_sparse_c
ategorical_accuracy: 0.9130 - 53s/epoch - 113ms/step
Epoch 29/100
469/469 - 53s - loss: 0.3935 - sparse_categorical_accuracy: 0.8724 - val_loss: 0.2711 - val_sparse_c
ategorical_accuracy: 0.9144 - 53s/epoch - 113ms/step
Epoch 30/100
469/469 - 53s - loss: 0.3908 - sparse_categorical_accuracy: 0.8716 - val_loss: 0.2632 - val_sparse_c
ategorical_accuracy: 0.9170 - 53s/epoch - 114ms/step
Epoch 31/100
```

```
469/469 - 53s - loss: 0.3878 - sparse_categorical_accuracy: 0.8735 - val_loss: 0.2658 - val_sparse_c
ategorical_accuracy: 0.9153 - 53s/epoch - 114ms/step
Epoch 32/100
469/469 - 53s - loss: 0.3882 - sparse_categorical_accuracy: 0.8737 - val_loss: 0.2604 - val_sparse_c
ategorical_accuracy: 0.9167 - 53s/epoch - 114ms/step
Epoch 33/100
469/469 - 53s - loss: 0.3811 - sparse_categorical_accuracy: 0.8761 - val_loss: 0.2533 - val_sparse_c
ategorical_accuracy: 0.9211 - 53s/epoch - 114ms/step
Epoch 34/100
469/469 - 53s - loss: 0.3848 - sparse_categorical_accuracy: 0.8737 - val_loss: 0.2555 - val_sparse_c
ategorical_accuracy: 0.9189 - 53s/epoch - 113ms/step
Epoch 35/100
469/469 - 53s - loss: 0.3778 - sparse_categorical_accuracy: 0.8762 - val_loss: 0.2534 - val_sparse_c
ategorical_accuracy: 0.9193 - 53s/epoch - 112ms/step
Epoch 36/100
469/469 - 53s - loss: 0.3716 - sparse_categorical_accuracy: 0.8789 - val_loss: 0.2492 - val_sparse_c
ategorical_accuracy: 0.9218 - 53s/epoch - 114ms/step
Epoch 37/100
469/469 - 53s - loss: 0.3721 - sparse_categorical_accuracy: 0.8789 - val_loss: 0.2484 - val_sparse_c
ategorical_accuracy: 0.9212 - 53s/epoch - 114ms/step
Epoch 38/100
469/469 - 53s - loss: 0.3710 - sparse_categorical_accuracy: 0.8796 - val_loss: 0.2442 - val_sparse_c
ategorical_accuracy: 0.9229 - 53s/epoch - 113ms/step
Epoch 39/100
469/469 - 53s - loss: 0.3662 - sparse_categorical_accuracy: 0.8809 - val_loss: 0.2373 - val_sparse_c
ategorical_accuracy: 0.9259 - 53s/epoch - 113ms/step
Epoch 40/100
469/469 - 53s - loss: 0.3593 - sparse_categorical_accuracy: 0.8839 - val_loss: 0.2393 - val_sparse_c
ategorical_accuracy: 0.9230 - 53s/epoch - 114ms/step
Epoch 41/100
469/469 - 53s - loss: 0.3586 - sparse_categorical_accuracy: 0.8838 - val_loss: 0.2359 - val_sparse_c
ategorical_accuracy: 0.9242 - 53s/epoch - 113ms/step
Epoch 42/100
469/469 - 54s - loss: 0.3509 - sparse_categorical_accuracy: 0.8872 - val_loss: 0.2356 - val_sparse_c
ategorical_accuracy: 0.9254 - 54s/epoch - 114ms/step
Epoch 43/100
469/469 - 54s - loss: 0.3441 - sparse_categorical_accuracy: 0.8891 - val_loss: 0.2309 - val_sparse_c
ategorical_accuracy: 0.9287 - 54s/epoch - 114ms/step
Epoch 44/100
469/469 - 53s - loss: 0.3358 - sparse_categorical_accuracy: 0.8913 - val_loss: 0.2327 - val_sparse_c
ategorical_accuracy: 0.9247 - 53s/epoch - 113ms/step
Epoch 45/100
469/469 - 53s - loss: 0.3334 - sparse_categorical_accuracy: 0.8919 - val_loss: 0.2203 - val_sparse_c
ategorical_accuracy: 0.9323 - 53s/epoch - 113ms/step
Epoch 46/100
469/469 - 53s - loss: 0.3231 - sparse_categorical_accuracy: 0.8963 - val_loss: 0.2160 - val_sparse_c
ategorical_accuracy: 0.9334 - 53s/epoch - 113ms/step
Epoch 47/100
469/469 - 53s - loss: 0.3105 - sparse_categorical_accuracy: 0.9007 - val_loss: 0.2097 - val_sparse_c
ategorical_accuracy: 0.9355 - 53s/epoch - 114ms/step
Epoch 48/100
469/469 - 53s - loss: 0.3008 - sparse_categorical_accuracy: 0.9026 - val_loss: 0.2008 - val_sparse_c
ategorical_accuracy: 0.9377 - 53s/epoch - 113ms/step
Epoch 49/100
469/469 - 53s - loss: 0.2933 - sparse_categorical_accuracy: 0.9046 - val_loss: 0.2002 - val_sparse_c
ategorical_accuracy: 0.9368 - 53s/epoch - 114ms/step
Epoch 50/100
469/469 - 53s - loss: 0.2813 - sparse_categorical_accuracy: 0.9102 - val_loss: 0.1913 - val_sparse_c
ategorical_accuracy: 0.9424 - 53s/epoch - 113ms/step
Epoch 51/100
469/469 - 53s - loss: 0.2705 - sparse_categorical_accuracy: 0.9140 - val_loss: 0.1869 - val_sparse_c
ategorical_accuracy: 0.9419 - 53s/epoch - 113ms/step
Epoch 52/100
469/469 - 53s - loss: 0.2678 - sparse_categorical_accuracy: 0.9151 - val_loss: 0.1839 - val_sparse_c
ategorical_accuracy: 0.9434 - 53s/epoch - 113ms/step
Epoch 53/100
469/469 - 53s - loss: 0.2629 - sparse_categorical_accuracy: 0.9157 - val_loss: 0.1849 - val_sparse_c
ategorical_accuracy: 0.9410 - 53s/epoch - 114ms/step
Epoch 54/100
469/469 - 53s - loss: 0.2551 - sparse_categorical_accuracy: 0.9191 - val_loss: 0.1746 - val_sparse_c
ategorical_accuracy: 0.9461 - 53s/epoch - 113ms/step
Epoch 55/100
469/469 - 53s - loss: 0.2547 - sparse_categorical_accuracy: 0.9189 - val_loss: 0.1732 - val_sparse_c
ategorical_accuracy: 0.9448 - 53s/epoch - 113ms/step
Epoch 56/100
469/469 - 53s - loss: 0.2527 - sparse_categorical_accuracy: 0.9204 - val_loss: 0.1726 - val_sparse_c
ategorical_accuracy: 0.9464 - 53s/epoch - 113ms/step
Epoch 57/100
469/469 - 53s - loss: 0.2455 - sparse_categorical_accuracy: 0.9212 - val_loss: 0.1645 - val_sparse_c
ategorical_accuracy: 0.9479 - 53s/epoch - 113ms/step
Epoch 58/100
469/469 - 53s - loss: 0.2431 - sparse_categorical_accuracy: 0.9231 - val_loss: 0.1632 - val_sparse_c
ategorical_accuracy: 0.9491 - 53s/epoch - 113ms/step
```

```
Epoch 59/100
469/469 - 53s - loss: 0.2376 - sparse_categorical_accuracy: 0.9252 - val_loss: 0.1637 - val_sparse_c
ategorical_accuracy: 0.9470 - 53s/epoch - 114ms/step
Epoch 60/100
469/469 - 53s - loss: 0.2367 - sparse_categorical_accuracy: 0.9247 - val_loss: 0.1689 - val_sparse_c
ategorical_accuracy: 0.9465 - 53s/epoch - 114ms/step
Epoch 61/100
469/469 - 53s - loss: 0.2320 - sparse_categorical_accuracy: 0.9269 - val_loss: 0.1601 - val_sparse_c
ategorical_accuracy: 0.9493 - 53s/epoch - 114ms/step
Epoch 62/100
469/469 - 53s - loss: 0.2325 - sparse_categorical_accuracy: 0.9262 - val_loss: 0.1660 - val_sparse_c
ategorical_accuracy: 0.9489 - 53s/epoch - 113ms/step
Epoch 63/100
469/469 - 53s - loss: 0.2261 - sparse_categorical_accuracy: 0.9283 - val_loss: 0.1548 - val_sparse_c
ategorical_accuracy: 0.9526 - 53s/epoch - 113ms/step
Epoch 64/100
469/469 - 53s - loss: 0.2244 - sparse_categorical_accuracy: 0.9289 - val_loss: 0.1536 - val_sparse_c
ategorical_accuracy: 0.9518 - 53s/epoch - 114ms/step
Epoch 65/100
469/469 - 53s - loss: 0.2206 - sparse_categorical_accuracy: 0.9293 - val_loss: 0.1485 - val_sparse_c
ategorical_accuracy: 0.9548 - 53s/epoch - 114ms/step
Epoch 66/100
469/469 - 54s - loss: 0.2189 - sparse_categorical_accuracy: 0.9309 - val_loss: 0.1476 - val_sparse_c
ategorical_accuracy: 0.9551 - 54s/epoch - 114ms/step
Epoch 67/100
469/469 - 53s - loss: 0.2149 - sparse_categorical_accuracy: 0.9315 - val_loss: 0.1466 - val_sparse_c
ategorical_accuracy: 0.9555 - 53s/epoch - 114ms/step
Epoch 68/100
469/469 - 53s - loss: 0.2131 - sparse_categorical_accuracy: 0.9322 - val_loss: 0.1465 - val_sparse_c
ategorical_accuracy: 0.9561 - 53s/epoch - 113ms/step
Epoch 69/100
469/469 - 53s - loss: 0.2122 - sparse_categorical_accuracy: 0.9328 - val_loss: 0.1441 - val_sparse_c
ategorical_accuracy: 0.9571 - 53s/epoch - 113ms/step
Epoch 70/100
469/469 - 53s - loss: 0.2110 - sparse_categorical_accuracy: 0.9335 - val_loss: 0.1453 - val_sparse_c
ategorical_accuracy: 0.9547 - 53s/epoch - 113ms/step
Epoch 71/100
469/469 - 53s - loss: 0.2064 - sparse_categorical_accuracy: 0.9348 - val_loss: 0.1447 - val_sparse_c
ategorical_accuracy: 0.9554 - 53s/epoch - 113ms/step
Epoch 72/100
469/469 - 53s - loss: 0.2046 - sparse_categorical_accuracy: 0.9348 - val_loss: 0.1365 - val_sparse_c
ategorical_accuracy: 0.9589 - 53s/epoch - 113ms/step
Epoch 73/100
469/469 - 53s - loss: 0.2034 - sparse_categorical_accuracy: 0.9351 - val_loss: 0.1347 - val_sparse_c
ategorical_accuracy: 0.9573 - 53s/epoch - 113ms/step
Epoch 74/100
469/469 - 53s - loss: 0.2018 - sparse_categorical_accuracy: 0.9351 - val_loss: 0.1358 - val_sparse_c
ategorical_accuracy: 0.9582 - 53s/epoch - 113ms/step
Epoch 75/100
469/469 - 53s - loss: 0.1982 - sparse_categorical_accuracy: 0.9377 - val_loss: 0.1338 - val_sparse_c
ategorical_accuracy: 0.9571 - 53s/epoch - 114ms/step
Epoch 76/100
469/469 - 53s - loss: 0.1963 - sparse_categorical_accuracy: 0.9381 - val_loss: 0.1309 - val_sparse_c
ategorical_accuracy: 0.9604 - 53s/epoch - 113ms/step
Epoch 77/100
469/469 - 53s - loss: 0.1959 - sparse_categorical_accuracy: 0.9384 - val_loss: 0.1331 - val_sparse_c
ategorical_accuracy: 0.9595 - 53s/epoch - 113ms/step
Epoch 78/100
469/469 - 53s - loss: 0.1929 - sparse_categorical_accuracy: 0.9395 - val_loss: 0.1293 - val_sparse_c
ategorical_accuracy: 0.9609 - 53s/epoch - 113ms/step
Epoch 79/100
469/469 - 53s - loss: 0.1926 - sparse_categorical_accuracy: 0.9400 - val_loss: 0.1261 - val_sparse_c
ategorical_accuracy: 0.9628 - 53s/epoch - 113ms/step
Epoch 80/100
469/469 - 53s - loss: 0.1907 - sparse_categorical_accuracy: 0.9396 - val_loss: 0.1258 - val_sparse_c
ategorical_accuracy: 0.9608 - 53s/epoch - 114ms/step
Epoch 81/100
469/469 - 54s - loss: 0.1902 - sparse_categorical_accuracy: 0.9387 - val_loss: 0.1310 - val_sparse_c
ategorical_accuracy: 0.9598 - 54s/epoch - 114ms/step
Epoch 82/100
469/469 - 53s - loss: 0.1888 - sparse_categorical_accuracy: 0.9395 - val_loss: 0.1252 - val_sparse_c
ategorical_accuracy: 0.9611 - 53s/epoch - 113ms/step
Epoch 83/100
469/469 - 53s - loss: 0.1900 - sparse_categorical_accuracy: 0.9391 - val_loss: 0.1241 - val_sparse_c
ategorical_accuracy: 0.9619 - 53s/epoch - 114ms/step
Epoch 84/100
469/469 - 53s - loss: 0.1863 - sparse_categorical_accuracy: 0.9404 - val_loss: 0.1236 - val_sparse_c
ategorical_accuracy: 0.9615 - 53s/epoch - 113ms/step
Epoch 85/100
469/469 - 53s - loss: 0.1833 - sparse_categorical_accuracy: 0.9415 - val_loss: 0.1209 - val_sparse_c
ategorical_accuracy: 0.9633 - 53s/epoch - 113ms/step
Epoch 86/100
469/469 - 53s - loss: 0.1855 - sparse_categorical_accuracy: 0.9407 - val_loss: 0.1203 - val_sparse_c
```

```
ategorical_accuracy: 0.9628 - 53s/epoch - 113ms/step
Epoch 87/100
469/469 - 53s - loss: 0.1846 - sparse_categorical_accuracy: 0.9419 - val_loss: 0.1198 - val_sparse_c
ategorical_accuracy: 0.9627 - 53s/epoch - 113ms/step
Epoch 88/100
469/469 - 53s - loss: 0.1830 - sparse_categorical_accuracy: 0.9416 - val_loss: 0.1200 - val_sparse_c
ategorical_accuracy: 0.9617 - 53s/epoch - 113ms/step
Epoch 89/100
469/469 - 54s - loss: 0.1795 - sparse_categorical_accuracy: 0.9419 - val_loss: 0.1192 - val_sparse_c
ategorical_accuracy: 0.9614 - 54s/epoch - 114ms/step
Epoch 90/100
469/469 - 54s - loss: 0.1776 - sparse_categorical_accuracy: 0.9436 - val_loss: 0.1207 - val_sparse_c
ategorical_accuracy: 0.9614 - 54s/epoch - 114ms/step
Epoch 91/100
469/469 - 53s - loss: 0.1791 - sparse_categorical_accuracy: 0.9427 - val_loss: 0.1197 - val_sparse_c
ategorical_accuracy: 0.9628 - 53s/epoch - 114ms/step
Epoch 92/100
469/469 - 53s - loss: 0.1769 - sparse_categorical_accuracy: 0.9433 - val_loss: 0.1149 - val_sparse_c
ategorical_accuracy: 0.9637 - 53s/epoch - 114ms/step
Epoch 93/100
469/469 - 53s - loss: 0.1778 - sparse_categorical_accuracy: 0.9428 - val_loss: 0.1161 - val_sparse_c
ategorical_accuracy: 0.9636 - 53s/epoch - 114ms/step
Epoch 94/100
469/469 - 53s - loss: 0.1794 - sparse_categorical_accuracy: 0.9430 - val_loss: 0.1169 - val_sparse_c
ategorical_accuracy: 0.9631 - 53s/epoch - 114ms/step
Epoch 95/100
469/469 - 53s - loss: 0.1750 - sparse_categorical_accuracy: 0.9431 - val_loss: 0.1160 - val_sparse_c
ategorical_accuracy: 0.9642 - 53s/epoch - 114ms/step
Epoch 96/100
469/469 - 53s - loss: 0.1738 - sparse_categorical_accuracy: 0.9451 - val_loss: 0.1135 - val_sparse_c
ategorical_accuracy: 0.9662 - 53s/epoch - 113ms/step
Epoch 97/100
469/469 - 53s - loss: 0.1726 - sparse_categorical_accuracy: 0.9447 - val_loss: 0.1170 - val_sparse_c
ategorical_accuracy: 0.9635 - 53s/epoch - 113ms/step
Epoch 98/100
469/469 - 53s - loss: 0.1747 - sparse_categorical_accuracy: 0.9437 - val_loss: 0.1136 - val_sparse_c
ategorical_accuracy: 0.9655 - 53s/epoch - 114ms/step
Epoch 99/100
469/469 - 53s - loss: 0.1716 - sparse_categorical_accuracy: 0.9450 - val_loss: 0.1117 - val_sparse_c
ategorical_accuracy: 0.9658 - 53s/epoch - 114ms/step
Epoch 100/100
469/469 - 54s - loss: 0.1733 - sparse_categorical_accuracy: 0.9442 - val_loss: 0.1127 - val_sparse_c
ategorical_accuracy: 0.9660 - 54s/epoch - 114ms/step
```

In [32]:

```
model6.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 4s - loss: 0.1127 - sparse_categorical_accuracy: 0.9660 - 4s/epoch - 14ms/step
```

Out[32]:

```
[0.11269041895866394, 0.9660000205039978]
```

Model7: `adam` optimizer with learning rate=e^{-3}, `random_normal` initializer, `dropout` regularization with rate=0.2.

In [33]:

```
input_layer = Input(shape=(28, 28,1))
x = inception_module(input_layer,
                     filters_1x1=1,
                     filters_3x3_reduce=1,
                     filters_3x3=1,
                     filters_5x5_reduce=1,
                     filters_5x5=1,
                     filters_pool_proj=1,
                     name='inception_3a')
x = tf.keras.layers.BatchNormalization()(x)
x = MaxPool2D((3, 3))(x)
x = tf.keras.layers.Flatten()(x)
x = Dropout(0.2)(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal')(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal')(x)
x = tf.keras.layers.BatchNormalization()(x)
x = Dense(10, activation='softmax', kernel_initializer='random_normal')(x)
model7 = Model(input_layer, x, name='inception_v1')
```

In [34]:

```
model7.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [35]:

```
history7 = model7.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 54s - loss: 0.8778 - sparse_categorical_accuracy: 0.7723 - val_loss: 0.2721 - val_sparse_c
ategorical_accuracy: 0.9356 - 54s/epoch - 115ms/step
Epoch 2/100
469/469 - 53s - loss: 0.2546 - sparse_categorical_accuracy: 0.9257 - val_loss: 0.1494 - val_sparse_c
ategorical_accuracy: 0.9580 - 53s/epoch - 112ms/step
Epoch 3/100
469/469 - 53s - loss: 0.1932 - sparse_categorical_accuracy: 0.9399 - val_loss: 0.1190 - val_sparse_c
ategorical_accuracy: 0.9653 - 53s/epoch - 113ms/step
Epoch 4/100
469/469 - 53s - loss: 0.1736 - sparse_categorical_accuracy: 0.9461 - val_loss: 0.1052 - val_sparse_c
ategorical_accuracy: 0.9692 - 53s/epoch - 113ms/step
Epoch 5/100
469/469 - 53s - loss: 0.1611 - sparse_categorical_accuracy: 0.9487 - val_loss: 0.1031 - val_sparse_c
ategorical_accuracy: 0.9684 - 53s/epoch - 113ms/step
Epoch 6/100
469/469 - 53s - loss: 0.1535 - sparse_categorical_accuracy: 0.9512 - val_loss: 0.0903 - val_sparse_c
ategorical_accuracy: 0.9736 - 53s/epoch - 113ms/step
Epoch 7/100
469/469 - 53s - loss: 0.1492 - sparse_categorical_accuracy: 0.9531 - val_loss: 0.0957 - val_sparse_c
ategorical_accuracy: 0.9702 - 53s/epoch - 113ms/step
Epoch 8/100
469/469 - 53s - loss: 0.1432 - sparse_categorical_accuracy: 0.9539 - val_loss: 0.0967 - val_sparse_c
ategorical_accuracy: 0.9698 - 53s/epoch - 113ms/step
Epoch 9/100
469/469 - 53s - loss: 0.1375 - sparse_categorical_accuracy: 0.9564 - val_loss: 0.0890 - val_sparse_c
ategorical_accuracy: 0.9724 - 53s/epoch - 113ms/step
Epoch 10/100
469/469 - 53s - loss: 0.1339 - sparse_categorical_accuracy: 0.9578 - val_loss: 0.0828 - val_sparse_c
ategorical_accuracy: 0.9736 - 53s/epoch - 114ms/step
Epoch 11/100
469/469 - 53s - loss: 0.1316 - sparse_categorical_accuracy: 0.9577 - val_loss: 0.0855 - val_sparse_c
ategorical_accuracy: 0.9735 - 53s/epoch - 114ms/step
Epoch 12/100
469/469 - 53s - loss: 0.1292 - sparse_categorical_accuracy: 0.9589 - val_loss: 0.0912 - val_sparse_c
ategorical_accuracy: 0.9713 - 53s/epoch - 113ms/step
Epoch 13/100
469/469 - 53s - loss: 0.1269 - sparse_categorical_accuracy: 0.9592 - val_loss: 0.0822 - val_sparse_c
ategorical_accuracy: 0.9751 - 53s/epoch - 113ms/step
Epoch 14/100
469/469 - 53s - loss: 0.1218 - sparse_categorical_accuracy: 0.9606 - val_loss: 0.0800 - val_sparse_c
ategorical_accuracy: 0.9764 - 53s/epoch - 114ms/step
Epoch 15/100
469/469 - 53s - loss: 0.1214 - sparse_categorical_accuracy: 0.9618 - val_loss: 0.0835 - val_sparse_c
ategorical_accuracy: 0.9737 - 53s/epoch - 113ms/step
Epoch 16/100
469/469 - 53s - loss: 0.1215 - sparse_categorical_accuracy: 0.9610 - val_loss: 0.0763 - val_sparse_c
ategorical_accuracy: 0.9757 - 53s/epoch - 113ms/step
Epoch 17/100
469/469 - 53s - loss: 0.1208 - sparse_categorical_accuracy: 0.9613 - val_loss: 0.0790 - val_sparse_c
ategorical_accuracy: 0.9755 - 53s/epoch - 113ms/step
Epoch 18/100
469/469 - 53s - loss: 0.1179 - sparse_categorical_accuracy: 0.9626 - val_loss: 0.0841 - val_sparse_c
ategorical_accuracy: 0.9731 - 53s/epoch - 113ms/step
Epoch 19/100
469/469 - 53s - loss: 0.1150 - sparse_categorical_accuracy: 0.9633 - val_loss: 0.0804 - val_sparse_c
ategorical_accuracy: 0.9774 - 53s/epoch - 113ms/step
Epoch 20/100
469/469 - 53s - loss: 0.1167 - sparse_categorical_accuracy: 0.9619 - val_loss: 0.0785 - val_sparse_c
ategorical_accuracy: 0.9763 - 53s/epoch - 113ms/step
Epoch 21/100
469/469 - 53s - loss: 0.1163 - sparse_categorical_accuracy: 0.9622 - val_loss: 0.0772 - val_sparse_c
ategorical_accuracy: 0.9761 - 53s/epoch - 113ms/step
Epoch 22/100
469/469 - 53s - loss: 0.1156 - sparse_categorical_accuracy: 0.9627 - val_loss: 0.0801 - val_sparse_c
ategorical_accuracy: 0.9747 - 53s/epoch - 113ms/step
Epoch 23/100
```

```
469/469 - 53s - loss: 0.1140 - sparse_categorical_accuracy: 0.9632 - val_loss: 0.0769 - val_sparse_c
ategorical_accuracy: 0.9760 - 53s/epoch - 113ms/step
Epoch 24/100
469/469 - 53s - loss: 0.1115 - sparse_categorical_accuracy: 0.9635 - val_loss: 0.0765 - val_sparse_c
ategorical_accuracy: 0.9772 - 53s/epoch - 112ms/step
Epoch 25/100
469/469 - 53s - loss: 0.1104 - sparse_categorical_accuracy: 0.9647 - val_loss: 0.0715 - val_sparse_c
ategorical_accuracy: 0.9785 - 53s/epoch - 113ms/step
Epoch 26/100
469/469 - 53s - loss: 0.1128 - sparse_categorical_accuracy: 0.9636 - val_loss: 0.0719 - val_sparse_c
ategorical_accuracy: 0.9778 - 53s/epoch - 113ms/step
Epoch 27/100
469/469 - 53s - loss: 0.1110 - sparse_categorical_accuracy: 0.9633 - val_loss: 0.0731 - val_sparse_c
ategorical_accuracy: 0.9769 - 53s/epoch - 113ms/step
Epoch 28/100
469/469 - 53s - loss: 0.1093 - sparse_categorical_accuracy: 0.9652 - val_loss: 0.0730 - val_sparse_c
ategorical_accuracy: 0.9775 - 53s/epoch - 113ms/step
Epoch 29/100
469/469 - 53s - loss: 0.1076 - sparse_categorical_accuracy: 0.9648 - val_loss: 0.0704 - val_sparse_c
ategorical_accuracy: 0.9777 - 53s/epoch - 113ms/step
Epoch 30/100
469/469 - 53s - loss: 0.1117 - sparse_categorical_accuracy: 0.9636 - val_loss: 0.0709 - val_sparse_c
ategorical_accuracy: 0.9793 - 53s/epoch - 112ms/step
Epoch 31/100
469/469 - 53s - loss: 0.1096 - sparse_categorical_accuracy: 0.9646 - val_loss: 0.0741 - val_sparse_c
ategorical_accuracy: 0.9764 - 53s/epoch - 113ms/step
Epoch 32/100
469/469 - 53s - loss: 0.1093 - sparse_categorical_accuracy: 0.9646 - val_loss: 0.0695 - val_sparse_c
ategorical_accuracy: 0.9777 - 53s/epoch - 113ms/step
Epoch 33/100
469/469 - 53s - loss: 0.1060 - sparse_categorical_accuracy: 0.9650 - val_loss: 0.0688 - val_sparse_c
ategorical_accuracy: 0.9798 - 53s/epoch - 113ms/step
Epoch 34/100
469/469 - 53s - loss: 0.1068 - sparse_categorical_accuracy: 0.9665 - val_loss: 0.0718 - val_sparse_c
ategorical_accuracy: 0.9775 - 53s/epoch - 112ms/step
Epoch 35/100
469/469 - 53s - loss: 0.1107 - sparse_categorical_accuracy: 0.9642 - val_loss: 0.0690 - val_sparse_c
ategorical_accuracy: 0.9787 - 53s/epoch - 112ms/step
Epoch 36/100
469/469 - 53s - loss: 0.1058 - sparse_categorical_accuracy: 0.9669 - val_loss: 0.0708 - val_sparse_c
ategorical_accuracy: 0.9786 - 53s/epoch - 113ms/step
Epoch 37/100
469/469 - 53s - loss: 0.1048 - sparse_categorical_accuracy: 0.9663 - val_loss: 0.0704 - val_sparse_c
ategorical_accuracy: 0.9779 - 53s/epoch - 113ms/step
Epoch 38/100
469/469 - 53s - loss: 0.1058 - sparse_categorical_accuracy: 0.9661 - val_loss: 0.0684 - val_sparse_c
ategorical_accuracy: 0.9794 - 53s/epoch - 113ms/step
Epoch 39/100
469/469 - 53s - loss: 0.1078 - sparse_categorical_accuracy: 0.9650 - val_loss: 0.0699 - val_sparse_c
ategorical_accuracy: 0.9786 - 53s/epoch - 113ms/step
Epoch 40/100
469/469 - 53s - loss: 0.1052 - sparse_categorical_accuracy: 0.9656 - val_loss: 0.0709 - val_sparse_c
ategorical_accuracy: 0.9778 - 53s/epoch - 113ms/step
Epoch 41/100
469/469 - 53s - loss: 0.1025 - sparse_categorical_accuracy: 0.9668 - val_loss: 0.0679 - val_sparse_c
ategorical_accuracy: 0.9790 - 53s/epoch - 113ms/step
Epoch 42/100
469/469 - 53s - loss: 0.1063 - sparse_categorical_accuracy: 0.9654 - val_loss: 0.0669 - val_sparse_c
ategorical_accuracy: 0.9785 - 53s/epoch - 113ms/step
Epoch 43/100
469/469 - 53s - loss: 0.1040 - sparse_categorical_accuracy: 0.9664 - val_loss: 0.0675 - val_sparse_c
ategorical_accuracy: 0.9787 - 53s/epoch - 113ms/step
Epoch 44/100
469/469 - 53s - loss: 0.1030 - sparse_categorical_accuracy: 0.9664 - val_loss: 0.0689 - val_sparse_c
ategorical_accuracy: 0.9784 - 53s/epoch - 113ms/step
Epoch 45/100
469/469 - 53s - loss: 0.1028 - sparse_categorical_accuracy: 0.9667 - val_loss: 0.0674 - val_sparse_c
ategorical_accuracy: 0.9788 - 53s/epoch - 113ms/step
Epoch 46/100
469/469 - 53s - loss: 0.1040 - sparse_categorical_accuracy: 0.9665 - val_loss: 0.0677 - val_sparse_c
ategorical_accuracy: 0.9779 - 53s/epoch - 113ms/step
Epoch 47/100
469/469 - 53s - loss: 0.1025 - sparse_categorical_accuracy: 0.9663 - val_loss: 0.0689 - val_sparse_c
ategorical_accuracy: 0.9810 - 53s/epoch - 113ms/step
Epoch 48/100
469/469 - 53s - loss: 0.1013 - sparse_categorical_accuracy: 0.9675 - val_loss: 0.0689 - val_sparse_c
ategorical_accuracy: 0.9794 - 53s/epoch - 114ms/step
Epoch 49/100
469/469 - 53s - loss: 0.1005 - sparse_categorical_accuracy: 0.9677 - val_loss: 0.0685 - val_sparse_c
ategorical_accuracy: 0.9787 - 53s/epoch - 113ms/step
Epoch 50/100
469/469 - 53s - loss: 0.1031 - sparse_categorical_accuracy: 0.9671 - val_loss: 0.0661 - val_sparse_c
ategorical_accuracy: 0.9799 - 53s/epoch - 113ms/step
```

```
Epoch 51/100
469/469 - 53s - loss: 0.1029 - sparse_categorical_accuracy: 0.9662 - val_loss: 0.0712 - val_sparse_c
ategorical_accuracy: 0.9783 - 53s/epoch - 113ms/step
Epoch 52/100
469/469 - 53s - loss: 0.1016 - sparse_categorical_accuracy: 0.9676 - val_loss: 0.0663 - val_sparse_c
ategorical_accuracy: 0.9792 - 53s/epoch - 114ms/step
Epoch 53/100
469/469 - 53s - loss: 0.1015 - sparse_categorical_accuracy: 0.9674 - val_loss: 0.0700 - val_sparse_c
ategorical_accuracy: 0.9781 - 53s/epoch - 113ms/step
Epoch 54/100
469/469 - 53s - loss: 0.1006 - sparse_categorical_accuracy: 0.9668 - val_loss: 0.0722 - val_sparse_c
ategorical_accuracy: 0.9769 - 53s/epoch - 113ms/step
Epoch 55/100
469/469 - 53s - loss: 0.0993 - sparse_categorical_accuracy: 0.9675 - val_loss: 0.0668 - val_sparse_c
ategorical_accuracy: 0.9797 - 53s/epoch - 113ms/step
Epoch 56/100
469/469 - 53s - loss: 0.0996 - sparse_categorical_accuracy: 0.9683 - val_loss: 0.0643 - val_sparse_c
ategorical_accuracy: 0.9811 - 53s/epoch - 113ms/step
Epoch 57/100
469/469 - 53s - loss: 0.1022 - sparse_categorical_accuracy: 0.9675 - val_loss: 0.0696 - val_sparse_c
ategorical_accuracy: 0.9792 - 53s/epoch - 113ms/step
Epoch 58/100
469/469 - 53s - loss: 0.0974 - sparse_categorical_accuracy: 0.9681 - val_loss: 0.0689 - val_sparse_c
ategorical_accuracy: 0.9796 - 53s/epoch - 113ms/step
Epoch 59/100
469/469 - 53s - loss: 0.0999 - sparse_categorical_accuracy: 0.9681 - val_loss: 0.0696 - val_sparse_c
ategorical_accuracy: 0.9795 - 53s/epoch - 114ms/step
Epoch 60/100
469/469 - 53s - loss: 0.1017 - sparse_categorical_accuracy: 0.9673 - val_loss: 0.0671 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 113ms/step
Epoch 61/100
469/469 - 53s - loss: 0.0994 - sparse_categorical_accuracy: 0.9674 - val_loss: 0.0662 - val_sparse_c
ategorical_accuracy: 0.9810 - 53s/epoch - 113ms/step
Epoch 62/100
469/469 - 53s - loss: 0.0983 - sparse_categorical_accuracy: 0.9681 - val_loss: 0.0685 - val_sparse_c
ategorical_accuracy: 0.9785 - 53s/epoch - 113ms/step
Epoch 63/100
469/469 - 53s - loss: 0.0972 - sparse_categorical_accuracy: 0.9685 - val_loss: 0.0732 - val_sparse_c
ategorical_accuracy: 0.9782 - 53s/epoch - 113ms/step
Epoch 64/100
469/469 - 53s - loss: 0.1002 - sparse_categorical_accuracy: 0.9671 - val_loss: 0.0652 - val_sparse_c
ategorical_accuracy: 0.9808 - 53s/epoch - 112ms/step
Epoch 65/100
469/469 - 53s - loss: 0.0990 - sparse_categorical_accuracy: 0.9673 - val_loss: 0.0707 - val_sparse_c
ategorical_accuracy: 0.9789 - 53s/epoch - 114ms/step
Epoch 66/100
469/469 - 53s - loss: 0.0983 - sparse_categorical_accuracy: 0.9678 - val_loss: 0.0655 - val_sparse_c
ategorical_accuracy: 0.9806 - 53s/epoch - 113ms/step
Epoch 67/100
469/469 - 53s - loss: 0.1000 - sparse_categorical_accuracy: 0.9673 - val_loss: 0.0653 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 113ms/step
Epoch 68/100
469/469 - 53s - loss: 0.0968 - sparse_categorical_accuracy: 0.9692 - val_loss: 0.0663 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 113ms/step
Epoch 69/100
469/469 - 53s - loss: 0.0979 - sparse_categorical_accuracy: 0.9675 - val_loss: 0.0682 - val_sparse_c
ategorical_accuracy: 0.9795 - 53s/epoch - 114ms/step
Epoch 70/100
469/469 - 53s - loss: 0.0972 - sparse_categorical_accuracy: 0.9687 - val_loss: 0.0650 - val_sparse_c
ategorical_accuracy: 0.9807 - 53s/epoch - 113ms/step
Epoch 71/100
469/469 - 53s - loss: 0.0966 - sparse_categorical_accuracy: 0.9687 - val_loss: 0.0649 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 114ms/step
Epoch 72/100
469/469 - 53s - loss: 0.0972 - sparse_categorical_accuracy: 0.9679 - val_loss: 0.0680 - val_sparse_c
ategorical_accuracy: 0.9785 - 53s/epoch - 114ms/step
Epoch 73/100
469/469 - 53s - loss: 0.0964 - sparse_categorical_accuracy: 0.9688 - val_loss: 0.0685 - val_sparse_c
ategorical_accuracy: 0.9784 - 53s/epoch - 113ms/step
Epoch 74/100
469/469 - 53s - loss: 0.0986 - sparse_categorical_accuracy: 0.9683 - val_loss: 0.0664 - val_sparse_c
ategorical_accuracy: 0.9788 - 53s/epoch - 114ms/step
Epoch 75/100
469/469 - 53s - loss: 0.0974 - sparse_categorical_accuracy: 0.9684 - val_loss: 0.0645 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 113ms/step
Epoch 76/100
469/469 - 53s - loss: 0.0948 - sparse_categorical_accuracy: 0.9688 - val_loss: 0.0656 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 114ms/step
Epoch 77/100
469/469 - 53s - loss: 0.0949 - sparse_categorical_accuracy: 0.9691 - val_loss: 0.0671 - val_sparse_c
ategorical_accuracy: 0.9794 - 53s/epoch - 113ms/step
Epoch 78/100
469/469 - 53s - loss: 0.0960 - sparse_categorical_accuracy: 0.9688 - val_loss: 0.0649 - val_sparse_c
```

```
ategorical_accuracy: 0.9804 - 53s/epoch - 113ms/step
Epoch 79/100
469/469 - 53s - loss: 0.0946 - sparse_categorical_accuracy: 0.9693 - val_loss: 0.0648 - val_sparse_c
ategorical_accuracy: 0.9804 - 53s/epoch - 114ms/step
Epoch 80/100
469/469 - 53s - loss: 0.0978 - sparse_categorical_accuracy: 0.9680 - val_loss: 0.0671 - val_sparse_c
ategorical_accuracy: 0.9800 - 53s/epoch - 113ms/step
Epoch 81/100
469/469 - 53s - loss: 0.0967 - sparse_categorical_accuracy: 0.9682 - val_loss: 0.0675 - val_sparse_c
ategorical_accuracy: 0.9785 - 53s/epoch - 113ms/step
Epoch 82/100
469/469 - 53s - loss: 0.0964 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0643 - val_sparse_c
ategorical_accuracy: 0.9806 - 53s/epoch - 114ms/step
Epoch 83/100
469/469 - 53s - loss: 0.0948 - sparse_categorical_accuracy: 0.9685 - val_loss: 0.0649 - val_sparse_c
ategorical_accuracy: 0.9796 - 53s/epoch - 114ms/step
Epoch 84/100
469/469 - 53s - loss: 0.0944 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0645 - val_sparse_c
ategorical_accuracy: 0.9801 - 53s/epoch - 113ms/step
Epoch 85/100
469/469 - 53s - loss: 0.0954 - sparse_categorical_accuracy: 0.9689 - val_loss: 0.0669 - val_sparse_c
ategorical_accuracy: 0.9797 - 53s/epoch - 113ms/step
Epoch 86/100
469/469 - 53s - loss: 0.0945 - sparse_categorical_accuracy: 0.9692 - val_loss: 0.0651 - val_sparse_c
ategorical_accuracy: 0.9805 - 53s/epoch - 113ms/step
Epoch 87/100
469/469 - 53s - loss: 0.0959 - sparse_categorical_accuracy: 0.9689 - val_loss: 0.0668 - val_sparse_c
ategorical_accuracy: 0.9789 - 53s/epoch - 114ms/step
Epoch 88/100
469/469 - 53s - loss: 0.0949 - sparse_categorical_accuracy: 0.9692 - val_loss: 0.0666 - val_sparse_c
ategorical_accuracy: 0.9800 - 53s/epoch - 113ms/step
Epoch 89/100
469/469 - 53s - loss: 0.0934 - sparse_categorical_accuracy: 0.9697 - val_loss: 0.0646 - val_sparse_c
ategorical_accuracy: 0.9794 - 53s/epoch - 113ms/step
Epoch 90/100
469/469 - 53s - loss: 0.0934 - sparse_categorical_accuracy: 0.9700 - val_loss: 0.0640 - val_sparse_c
ategorical_accuracy: 0.9794 - 53s/epoch - 112ms/step
Epoch 91/100
469/469 - 53s - loss: 0.0945 - sparse_categorical_accuracy: 0.9694 - val_loss: 0.0641 - val_sparse_c
ategorical_accuracy: 0.9808 - 53s/epoch - 112ms/step
Epoch 92/100
469/469 - 53s - loss: 0.0936 - sparse_categorical_accuracy: 0.9691 - val_loss: 0.0663 - val_sparse_c
ategorical_accuracy: 0.9796 - 53s/epoch - 113ms/step
Epoch 93/100
469/469 - 53s - loss: 0.0925 - sparse_categorical_accuracy: 0.9701 - val_loss: 0.0657 - val_sparse_c
ategorical_accuracy: 0.9785 - 53s/epoch - 114ms/step
Epoch 94/100
469/469 - 53s - loss: 0.0929 - sparse_categorical_accuracy: 0.9698 - val_loss: 0.0657 - val_sparse_c
ategorical_accuracy: 0.9803 - 53s/epoch - 113ms/step
Epoch 95/100
469/469 - 53s - loss: 0.0932 - sparse_categorical_accuracy: 0.9696 - val_loss: 0.0679 - val_sparse_c
ategorical_accuracy: 0.9793 - 53s/epoch - 113ms/step
Epoch 96/100
469/469 - 53s - loss: 0.0913 - sparse_categorical_accuracy: 0.9696 - val_loss: 0.0679 - val_sparse_c
ategorical_accuracy: 0.9792 - 53s/epoch - 113ms/step
Epoch 97/100
469/469 - 53s - loss: 0.0940 - sparse_categorical_accuracy: 0.9690 - val_loss: 0.0672 - val_sparse_c
ategorical_accuracy: 0.9790 - 53s/epoch - 113ms/step
Epoch 98/100
469/469 - 53s - loss: 0.0925 - sparse_categorical_accuracy: 0.9701 - val_loss: 0.0650 - val_sparse_c
ategorical_accuracy: 0.9795 - 53s/epoch - 114ms/step
Epoch 99/100
469/469 - 53s - loss: 0.0940 - sparse_categorical_accuracy: 0.9695 - val_loss: 0.0654 - val_sparse_c
ategorical_accuracy: 0.9806 - 53s/epoch - 113ms/step
Epoch 100/100
469/469 - 53s - loss: 0.0931 - sparse_categorical_accuracy: 0.9700 - val_loss: 0.0707 - val_sparse_c
ategorical_accuracy: 0.9778 - 53s/epoch - 113ms/step
```

In [36]:

```
model7.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 4s - loss: 0.0707 - sparse_categorical_accuracy: 0.9778 - 4s/epoch - 13ms/step
```

Out[36]:

```
[0.07072369754314423, 0.9778000116348267]
```

Model8: `adam` optimizer with learning rate=e^{-4}, `random_normal` initializer, `dropout` regularization with rate=0.2.

In [37]:

```
input_layer = Input(shape=(28, 28,1))
x = inception_module(input_layer,
                     filters_1x1=1,
                     filters_3x3_reduce=1,
                     filters_3x3=1,
                     filters_5x5_reduce=1,
                     filters_5x5=1,
                     filters_pool_proj=1,
                     name='inception_3a')
x = tf.keras.layers.BatchNormalization()(x)
x = MaxPool2D((3, 3))(x)
x = tf.keras.layers.Flatten()(x)
x = Dropout(0.2)(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal')(x)
x = tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal')(x)
x = tf.keras.layers.BatchNormalization()(x)
x = Dense(10, activation='softmax', kernel_initializer='random_normal')(x)
model8 = Model(input_layer, x, name='inception_v1')
```

In [38]:

```
model8.compile(optimizer=tf.keras.optimizers.Adam(0.0001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [39]:

```
history8 = model8.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 54s - loss: 2.0019 - sparse_categorical_accuracy: 0.3940 - val_loss: 1.6929 - val_sparse_c
ategorical_accuracy: 0.6367 - 54s/epoch - 116ms/step
Epoch 2/100
469/469 - 53s - loss: 1.5047 - sparse_categorical_accuracy: 0.6314 - val_loss: 1.2433 - val_sparse_c
ategorical_accuracy: 0.7304 - 53s/epoch - 113ms/step
Epoch 3/100
469/469 - 53s - loss: 1.1574 - sparse_categorical_accuracy: 0.7271 - val_loss: 0.9346 - val_sparse_c
ategorical_accuracy: 0.8106 - 53s/epoch - 114ms/step
Epoch 4/100
469/469 - 53s - loss: 0.9142 - sparse_categorical_accuracy: 0.7806 - val_loss: 0.7155 - val_sparse_c
ategorical_accuracy: 0.8499 - 53s/epoch - 114ms/step
Epoch 5/100
469/469 - 53s - loss: 0.7577 - sparse_categorical_accuracy: 0.8063 - val_loss: 0.5890 - val_sparse_c
ategorical_accuracy: 0.8720 - 53s/epoch - 113ms/step
Epoch 6/100
469/469 - 53s - loss: 0.6601 - sparse_categorical_accuracy: 0.8213 - val_loss: 0.5017 - val_sparse_c
ategorical_accuracy: 0.8784 - 53s/epoch - 113ms/step
Epoch 7/100
469/469 - 53s - loss: 0.5968 - sparse_categorical_accuracy: 0.8299 - val_loss: 0.4491 - val_sparse_c
ategorical_accuracy: 0.8827 - 53s/epoch - 113ms/step
Epoch 8/100
469/469 - 53s - loss: 0.5527 - sparse_categorical_accuracy: 0.8362 - val_loss: 0.4164 - val_sparse_c
ategorical_accuracy: 0.8863 - 53s/epoch - 113ms/step
Epoch 9/100
469/469 - 53s - loss: 0.5238 - sparse_categorical_accuracy: 0.8412 - val_loss: 0.3873 - val_sparse_c
ategorical_accuracy: 0.8911 - 53s/epoch - 114ms/step
Epoch 10/100
469/469 - 53s - loss: 0.5042 - sparse_categorical_accuracy: 0.8446 - val_loss: 0.3692 - val_sparse_c
ategorical_accuracy: 0.8924 - 53s/epoch - 113ms/step
Epoch 11/100
469/469 - 53s - loss: 0.4899 - sparse_categorical_accuracy: 0.8471 - val_loss: 0.3590 - val_sparse_c
ategorical_accuracy: 0.8952 - 53s/epoch - 114ms/step
Epoch 12/100
469/469 - 53s - loss: 0.4757 - sparse_categorical_accuracy: 0.8503 - val_loss: 0.3461 - val_sparse_c
ategorical_accuracy: 0.8972 - 53s/epoch - 113ms/step
Epoch 13/100
469/469 - 53s - loss: 0.4653 - sparse_categorical_accuracy: 0.8526 - val_loss: 0.3398 - val_sparse_c
ategorical_accuracy: 0.8956 - 53s/epoch - 113ms/step
Epoch 14/100
469/469 - 53s - loss: 0.4582 - sparse_categorical_accuracy: 0.8534 - val_loss: 0.3286 - val_sparse_c
ategorical_accuracy: 0.8993 - 53s/epoch - 114ms/step
Epoch 15/100
469/469 - 53s - loss: 0.4462 - sparse_categorical_accuracy: 0.8582 - val_loss: 0.3223 - val_sparse_c
ategorical_accuracy: 0.9027 - 53s/epoch - 113ms/step
Epoch 16/100
```

```
469/469 - 53s - loss: 0.4403 - sparse_categorical_accuracy: 0.8585 - val_loss: 0.3140 - val_sparse_c
ategorical_accuracy: 0.9044 - 53s/epoch - 113ms/step
Epoch 17/100
469/469 - 53s - loss: 0.4351 - sparse_categorical_accuracy: 0.8611 - val_loss: 0.3163 - val_sparse_c
ategorical_accuracy: 0.9031 - 53s/epoch - 113ms/step
Epoch 18/100
469/469 - 53s - loss: 0.4292 - sparse_categorical_accuracy: 0.8606 - val_loss: 0.3074 - val_sparse_c
ategorical_accuracy: 0.9054 - 53s/epoch - 113ms/step
Epoch 19/100
469/469 - 53s - loss: 0.4244 - sparse_categorical_accuracy: 0.8620 - val_loss: 0.3124 - val_sparse_c
ategorical_accuracy: 0.9026 - 53s/epoch - 113ms/step
Epoch 20/100
469/469 - 53s - loss: 0.4197 - sparse_categorical_accuracy: 0.8638 - val_loss: 0.3026 - val_sparse_c
ategorical_accuracy: 0.9077 - 53s/epoch - 113ms/step
Epoch 21/100
469/469 - 53s - loss: 0.4108 - sparse_categorical_accuracy: 0.8677 - val_loss: 0.2987 - val_sparse_c
ategorical_accuracy: 0.9065 - 53s/epoch - 113ms/step
Epoch 22/100
469/469 - 53s - loss: 0.4063 - sparse_categorical_accuracy: 0.8676 - val_loss: 0.2945 - val_sparse_c
ategorical_accuracy: 0.9088 - 53s/epoch - 114ms/step
Epoch 23/100
469/469 - 53s - loss: 0.4037 - sparse_categorical_accuracy: 0.8678 - val_loss: 0.2865 - val_sparse_c
ategorical_accuracy: 0.9108 - 53s/epoch - 114ms/step
Epoch 24/100
469/469 - 53s - loss: 0.4024 - sparse_categorical_accuracy: 0.8700 - val_loss: 0.2835 - val_sparse_c
ategorical_accuracy: 0.9128 - 53s/epoch - 113ms/step
Epoch 25/100
469/469 - 53s - loss: 0.3940 - sparse_categorical_accuracy: 0.8708 - val_loss: 0.2794 - val_sparse_c
ategorical_accuracy: 0.9138 - 53s/epoch - 113ms/step
Epoch 26/100
469/469 - 53s - loss: 0.3888 - sparse_categorical_accuracy: 0.8741 - val_loss: 0.2770 - val_sparse_c
ategorical_accuracy: 0.9169 - 53s/epoch - 113ms/step
Epoch 27/100
469/469 - 53s - loss: 0.3830 - sparse_categorical_accuracy: 0.8743 - val_loss: 0.2718 - val_sparse_c
ategorical_accuracy: 0.9157 - 53s/epoch - 114ms/step
Epoch 28/100
469/469 - 53s - loss: 0.3741 - sparse_categorical_accuracy: 0.8792 - val_loss: 0.2672 - val_sparse_c
ategorical_accuracy: 0.9190 - 53s/epoch - 113ms/step
Epoch 29/100
469/469 - 53s - loss: 0.3681 - sparse_categorical_accuracy: 0.8800 - val_loss: 0.2636 - val_sparse_c
ategorical_accuracy: 0.9198 - 53s/epoch - 114ms/step
Epoch 30/100
469/469 - 53s - loss: 0.3621 - sparse_categorical_accuracy: 0.8824 - val_loss: 0.2565 - val_sparse_c
ategorical_accuracy: 0.9205 - 53s/epoch - 113ms/step
Epoch 31/100
469/469 - 53s - loss: 0.3537 - sparse_categorical_accuracy: 0.8849 - val_loss: 0.2528 - val_sparse_c
ategorical_accuracy: 0.9223 - 53s/epoch - 113ms/step
Epoch 32/100
469/469 - 53s - loss: 0.3489 - sparse_categorical_accuracy: 0.8870 - val_loss: 0.2499 - val_sparse_c
ategorical_accuracy: 0.9239 - 53s/epoch - 113ms/step
Epoch 33/100
469/469 - 53s - loss: 0.3386 - sparse_categorical_accuracy: 0.8911 - val_loss: 0.2460 - val_sparse_c
ategorical_accuracy: 0.9252 - 53s/epoch - 114ms/step
Epoch 34/100
469/469 - 53s - loss: 0.3356 - sparse_categorical_accuracy: 0.8909 - val_loss: 0.2432 - val_sparse_c
ategorical_accuracy: 0.9252 - 53s/epoch - 113ms/step
Epoch 35/100
469/469 - 53s - loss: 0.3275 - sparse_categorical_accuracy: 0.8936 - val_loss: 0.2377 - val_sparse_c
ategorical_accuracy: 0.9264 - 53s/epoch - 113ms/step
Epoch 36/100
469/469 - 53s - loss: 0.3235 - sparse_categorical_accuracy: 0.8949 - val_loss: 0.2344 - val_sparse_c
ategorical_accuracy: 0.9272 - 53s/epoch - 113ms/step
Epoch 37/100
469/469 - 53s - loss: 0.3204 - sparse_categorical_accuracy: 0.8957 - val_loss: 0.2310 - val_sparse_c
ategorical_accuracy: 0.9296 - 53s/epoch - 114ms/step
Epoch 38/100
469/469 - 53s - loss: 0.3098 - sparse_categorical_accuracy: 0.8995 - val_loss: 0.2257 - val_sparse_c
ategorical_accuracy: 0.9321 - 53s/epoch - 114ms/step
Epoch 39/100
469/469 - 53s - loss: 0.3116 - sparse_categorical_accuracy: 0.8997 - val_loss: 0.2261 - val_sparse_c
ategorical_accuracy: 0.9319 - 53s/epoch - 114ms/step
Epoch 40/100
469/469 - 53s - loss: 0.3069 - sparse_categorical_accuracy: 0.9010 - val_loss: 0.2219 - val_sparse_c
ategorical_accuracy: 0.9327 - 53s/epoch - 113ms/step
Epoch 41/100
469/469 - 53s - loss: 0.3031 - sparse_categorical_accuracy: 0.9024 - val_loss: 0.2197 - val_sparse_c
ategorical_accuracy: 0.9330 - 53s/epoch - 113ms/step
Epoch 42/100
469/469 - 53s - loss: 0.3013 - sparse_categorical_accuracy: 0.9024 - val_loss: 0.2191 - val_sparse_c
ategorical_accuracy: 0.9337 - 53s/epoch - 114ms/step
Epoch 43/100
469/469 - 53s - loss: 0.2980 - sparse_categorical_accuracy: 0.9038 - val_loss: 0.2184 - val_sparse_c
ategorical_accuracy: 0.9329 - 53s/epoch - 113ms/step
```

```
Epoch 44/100
469/469 - 53s - loss: 0.2953 - sparse_categorical_accuracy: 0.9047 - val_loss: 0.2124 - val_sparse_c
ategorical_accuracy: 0.9352 - 53s/epoch - 114ms/step
Epoch 45/100
469/469 - 53s - loss: 0.2940 - sparse_categorical_accuracy: 0.9045 - val_loss: 0.2171 - val_sparse_c
ategorical_accuracy: 0.9334 - 53s/epoch - 114ms/step
Epoch 46/100
469/469 - 53s - loss: 0.2939 - sparse_categorical_accuracy: 0.9051 - val_loss: 0.2138 - val_sparse_c
ategorical_accuracy: 0.9352 - 53s/epoch - 113ms/step
Epoch 47/100
469/469 - 53s - loss: 0.2906 - sparse_categorical_accuracy: 0.9049 - val_loss: 0.2093 - val_sparse_c
ategorical_accuracy: 0.9363 - 53s/epoch - 113ms/step
Epoch 48/100
469/469 - 53s - loss: 0.2878 - sparse_categorical_accuracy: 0.9077 - val_loss: 0.2077 - val_sparse_c
ategorical_accuracy: 0.9369 - 53s/epoch - 113ms/step
Epoch 49/100
469/469 - 53s - loss: 0.2858 - sparse_categorical_accuracy: 0.9081 - val_loss: 0.2086 - val_sparse_c
ategorical_accuracy: 0.9368 - 53s/epoch - 113ms/step
Epoch 50/100
469/469 - 53s - loss: 0.2842 - sparse_categorical_accuracy: 0.9076 - val_loss: 0.2053 - val_sparse_c
ategorical_accuracy: 0.9368 - 53s/epoch - 113ms/step
Epoch 51/100
469/469 - 53s - loss: 0.2842 - sparse_categorical_accuracy: 0.9089 - val_loss: 0.2052 - val_sparse_c
ategorical_accuracy: 0.9390 - 53s/epoch - 113ms/step
Epoch 52/100
469/469 - 53s - loss: 0.2819 - sparse_categorical_accuracy: 0.9086 - val_loss: 0.2024 - val_sparse_c
ategorical_accuracy: 0.9394 - 53s/epoch - 113ms/step
Epoch 53/100
469/469 - 53s - loss: 0.2808 - sparse_categorical_accuracy: 0.9090 - val_loss: 0.2050 - val_sparse_c
ategorical_accuracy: 0.9388 - 53s/epoch - 113ms/step
Epoch 54/100
469/469 - 53s - loss: 0.2772 - sparse_categorical_accuracy: 0.9109 - val_loss: 0.2021 - val_sparse_c
ategorical_accuracy: 0.9397 - 53s/epoch - 114ms/step
Epoch 55/100
469/469 - 53s - loss: 0.2781 - sparse_categorical_accuracy: 0.9103 - val_loss: 0.1999 - val_sparse_c
ategorical_accuracy: 0.9413 - 53s/epoch - 114ms/step
Epoch 56/100
469/469 - 53s - loss: 0.2746 - sparse_categorical_accuracy: 0.9118 - val_loss: 0.1987 - val_sparse_c
ategorical_accuracy: 0.9399 - 53s/epoch - 113ms/step
Epoch 57/100
469/469 - 53s - loss: 0.2737 - sparse_categorical_accuracy: 0.9110 - val_loss: 0.2010 - val_sparse_c
ategorical_accuracy: 0.9389 - 53s/epoch - 114ms/step
Epoch 58/100
469/469 - 53s - loss: 0.2757 - sparse_categorical_accuracy: 0.9115 - val_loss: 0.1956 - val_sparse_c
ategorical_accuracy: 0.9412 - 53s/epoch - 113ms/step
Epoch 59/100
469/469 - 53s - loss: 0.2732 - sparse_categorical_accuracy: 0.9115 - val_loss: 0.1949 - val_sparse_c
ategorical_accuracy: 0.9415 - 53s/epoch - 113ms/step
Epoch 60/100
469/469 - 53s - loss: 0.2735 - sparse_categorical_accuracy: 0.9115 - val_loss: 0.1960 - val_sparse_c
ategorical_accuracy: 0.9401 - 53s/epoch - 113ms/step
Epoch 61/100
469/469 - 53s - loss: 0.2667 - sparse_categorical_accuracy: 0.9128 - val_loss: 0.1933 - val_sparse_c
ategorical_accuracy: 0.9429 - 53s/epoch - 113ms/step
Epoch 62/100
469/469 - 53s - loss: 0.2702 - sparse_categorical_accuracy: 0.9134 - val_loss: 0.1925 - val_sparse_c
ategorical_accuracy: 0.9422 - 53s/epoch - 114ms/step
Epoch 63/100
469/469 - 53s - loss: 0.2688 - sparse_categorical_accuracy: 0.9130 - val_loss: 0.1924 - val_sparse_c
ategorical_accuracy: 0.9430 - 53s/epoch - 114ms/step
Epoch 64/100
469/469 - 53s - loss: 0.2687 - sparse_categorical_accuracy: 0.9124 - val_loss: 0.1909 - val_sparse_c
ategorical_accuracy: 0.9431 - 53s/epoch - 114ms/step
Epoch 65/100
469/469 - 53s - loss: 0.2679 - sparse_categorical_accuracy: 0.9128 - val_loss: 0.1916 - val_sparse_c
ategorical_accuracy: 0.9430 - 53s/epoch - 114ms/step
Epoch 66/100
469/469 - 53s - loss: 0.2645 - sparse_categorical_accuracy: 0.9160 - val_loss: 0.1901 - val_sparse_c
ategorical_accuracy: 0.9436 - 53s/epoch - 114ms/step
Epoch 67/100
469/469 - 53s - loss: 0.2627 - sparse_categorical_accuracy: 0.9160 - val_loss: 0.1876 - val_sparse_c
ategorical_accuracy: 0.9441 - 53s/epoch - 114ms/step
Epoch 68/100
469/469 - 53s - loss: 0.2641 - sparse_categorical_accuracy: 0.9140 - val_loss: 0.1906 - val_sparse_c
ategorical_accuracy: 0.9429 - 53s/epoch - 114ms/step
Epoch 69/100
469/469 - 53s - loss: 0.2618 - sparse_categorical_accuracy: 0.9163 - val_loss: 0.1878 - val_sparse_c
ategorical_accuracy: 0.9434 - 53s/epoch - 113ms/step
Epoch 70/100
469/469 - 53s - loss: 0.2609 - sparse_categorical_accuracy: 0.9159 - val_loss: 0.1879 - val_sparse_c
ategorical_accuracy: 0.9435 - 53s/epoch - 114ms/step
Epoch 71/100
469/469 - 54s - loss: 0.2590 - sparse_categorical_accuracy: 0.9171 - val_loss: 0.1870 - val_sparse_c
```

```
ategorical_accuracy: 0.9445 - 54s/epoch - 114ms/step
Epoch 72/100
469/469 - 53s - loss: 0.2593 - sparse_categorical_accuracy: 0.9177 - val_loss: 0.1877 - val_sparse_c
ategorical_accuracy: 0.9433 - 53s/epoch - 114ms/step
Epoch 73/100
469/469 - 54s - loss: 0.2597 - sparse_categorical_accuracy: 0.9153 - val_loss: 0.1865 - val_sparse_c
ategorical_accuracy: 0.9438 - 54s/epoch - 114ms/step
Epoch 74/100
469/469 - 53s - loss: 0.2567 - sparse_categorical_accuracy: 0.9169 - val_loss: 0.1833 - val_sparse_c
ategorical_accuracy: 0.9450 - 53s/epoch - 114ms/step
Epoch 75/100
469/469 - 53s - loss: 0.2576 - sparse_categorical_accuracy: 0.9173 - val_loss: 0.1832 - val_sparse_c
ategorical_accuracy: 0.9455 - 53s/epoch - 113ms/step
Epoch 76/100
469/469 - 53s - loss: 0.2564 - sparse_categorical_accuracy: 0.9162 - val_loss: 0.1828 - val_sparse_c
ategorical_accuracy: 0.9453 - 53s/epoch - 114ms/step
Epoch 77/100
469/469 - 53s - loss: 0.2551 - sparse_categorical_accuracy: 0.9181 - val_loss: 0.1828 - val_sparse_c
ategorical_accuracy: 0.9449 - 53s/epoch - 114ms/step
Epoch 78/100
469/469 - 53s - loss: 0.2529 - sparse_categorical_accuracy: 0.9185 - val_loss: 0.1820 - val_sparse_c
ategorical_accuracy: 0.9460 - 53s/epoch - 114ms/step
Epoch 79/100
469/469 - 53s - loss: 0.2517 - sparse_categorical_accuracy: 0.9191 - val_loss: 0.1816 - val_sparse_c
ategorical_accuracy: 0.9446 - 53s/epoch - 114ms/step
Epoch 80/100
469/469 - 53s - loss: 0.2513 - sparse_categorical_accuracy: 0.9197 - val_loss: 0.1808 - val_sparse_c
ategorical_accuracy: 0.9452 - 53s/epoch - 114ms/step
Epoch 81/100
469/469 - 53s - loss: 0.2521 - sparse_categorical_accuracy: 0.9198 - val_loss: 0.1795 - val_sparse_c
ategorical_accuracy: 0.9442 - 53s/epoch - 113ms/step
Epoch 82/100
469/469 - 53s - loss: 0.2510 - sparse_categorical_accuracy: 0.9193 - val_loss: 0.1795 - val_sparse_c
ategorical_accuracy: 0.9458 - 53s/epoch - 114ms/step
Epoch 83/100
469/469 - 53s - loss: 0.2501 - sparse_categorical_accuracy: 0.9195 - val_loss: 0.1778 - val_sparse_c
ategorical_accuracy: 0.9466 - 53s/epoch - 114ms/step
Epoch 84/100
469/469 - 53s - loss: 0.2494 - sparse_categorical_accuracy: 0.9195 - val_loss: 0.1760 - val_sparse_c
ategorical_accuracy: 0.9470 - 53s/epoch - 114ms/step
Epoch 85/100
469/469 - 53s - loss: 0.2481 - sparse_categorical_accuracy: 0.9189 - val_loss: 0.1792 - val_sparse_c
ategorical_accuracy: 0.9448 - 53s/epoch - 113ms/step
Epoch 86/100
469/469 - 53s - loss: 0.2474 - sparse_categorical_accuracy: 0.9199 - val_loss: 0.1733 - val_sparse_c
ategorical_accuracy: 0.9475 - 53s/epoch - 113ms/step
Epoch 87/100
469/469 - 53s - loss: 0.2475 - sparse_categorical_accuracy: 0.9208 - val_loss: 0.1737 - val_sparse_c
ategorical_accuracy: 0.9484 - 53s/epoch - 113ms/step
Epoch 88/100
469/469 - 53s - loss: 0.2471 - sparse_categorical_accuracy: 0.9199 - val_loss: 0.1756 - val_sparse_c
ategorical_accuracy: 0.9470 - 53s/epoch - 113ms/step
Epoch 89/100
469/469 - 53s - loss: 0.2452 - sparse_categorical_accuracy: 0.9196 - val_loss: 0.1747 - val_sparse_c
ategorical_accuracy: 0.9469 - 53s/epoch - 113ms/step
Epoch 90/100
469/469 - 53s - loss: 0.2430 - sparse_categorical_accuracy: 0.9212 - val_loss: 0.1726 - val_sparse_c
ategorical_accuracy: 0.9481 - 53s/epoch - 113ms/step
Epoch 91/100
469/469 - 53s - loss: 0.2389 - sparse_categorical_accuracy: 0.9233 - val_loss: 0.1715 - val_sparse_c
ategorical_accuracy: 0.9493 - 53s/epoch - 113ms/step
Epoch 92/100
469/469 - 53s - loss: 0.2429 - sparse_categorical_accuracy: 0.9215 - val_loss: 0.1721 - val_sparse_c
ategorical_accuracy: 0.9482 - 53s/epoch - 113ms/step
Epoch 93/100
469/469 - 53s - loss: 0.2428 - sparse_categorical_accuracy: 0.9219 - val_loss: 0.1706 - val_sparse_c
ategorical_accuracy: 0.9485 - 53s/epoch - 114ms/step
Epoch 94/100
469/469 - 53s - loss: 0.2397 - sparse_categorical_accuracy: 0.9218 - val_loss: 0.1706 - val_sparse_c
ategorical_accuracy: 0.9486 - 53s/epoch - 113ms/step
Epoch 95/100
469/469 - 53s - loss: 0.2389 - sparse_categorical_accuracy: 0.9228 - val_loss: 0.1702 - val_sparse_c
ategorical_accuracy: 0.9481 - 53s/epoch - 114ms/step
Epoch 96/100
469/469 - 53s - loss: 0.2347 - sparse_categorical_accuracy: 0.9252 - val_loss: 0.1674 - val_sparse_c
ategorical_accuracy: 0.9498 - 53s/epoch - 114ms/step
Epoch 97/100
469/469 - 54s - loss: 0.2360 - sparse_categorical_accuracy: 0.9250 - val_loss: 0.1674 - val_sparse_c
ategorical_accuracy: 0.9495 - 54s/epoch - 114ms/step
Epoch 98/100
469/469 - 53s - loss: 0.2364 - sparse_categorical_accuracy: 0.9243 - val_loss: 0.1663 - val_sparse_c
ategorical_accuracy: 0.9496 - 53s/epoch - 113ms/step
Epoch 99/100
```

```
469/469 - 53s - loss: 0.2346 - sparse_categorical_accuracy: 0.9250 - val_loss: 0.1645 - val_sparse_c
ategorical_accuracy: 0.9505 - 53s/epoch - 114ms/step
Epoch 100/100
469/469 - 53s - loss: 0.2301 - sparse_categorical_accuracy: 0.9252 - val_loss: 0.1641 - val_sparse_c
ategorical_accuracy: 0.9500 - 53s/epoch - 114ms/step
```

In [40]:

```
model8.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 4s - loss: 0.1641 - sparse_categorical_accuracy: 0.9500 - 4s/epoch - 13ms/step
```

Out[40]:

```
[0.1640833020210266, 0.949999988079071]
```

Plots

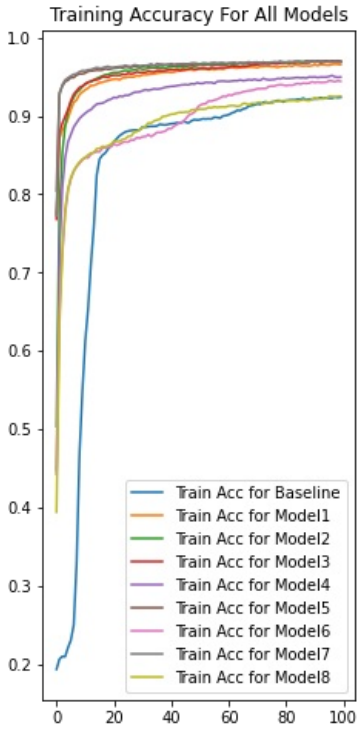
In [41]:

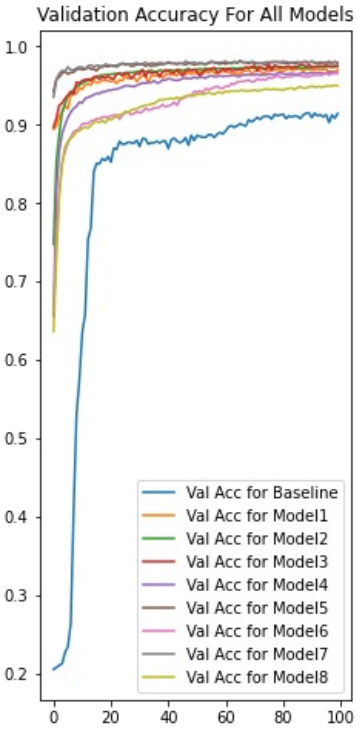
```
import matplotlib.pyplot as plt
```

Plots for different kernel initializers

In [42]:

```
training_accuracy0 = history.history['sparse_categorical_accuracy']
validation_accuracy0 = history.history['val_sparse_categorical_accuracy']

training_accuracy1 = history1.history['sparse_categorical_accuracy']
validation_accuracy1 = history1.history['val_sparse_categorical_accuracy']

training_accuracy2 = history2.history['sparse_categorical_accuracy']
validation_accuracy2 = history2.history['val_sparse_categorical_accuracy']

training_accuracy3 = history3.history['sparse_categorical_accuracy']
validation_accuracy3 = history3.history['val_sparse_categorical_accuracy']

training_accuracy4 = history4.history['sparse_categorical_accuracy']
validation_accuracy4 = history4.history['val_sparse_categorical_accuracy']

training_accuracy5 = history5.history['sparse_categorical_accuracy']
validation_accuracy5 = history5.history['val_sparse_categorical_accuracy']

training_accuracy6 = history6.history['sparse_categorical_accuracy']
validation_accuracy6 = history6.history['val_sparse_categorical_accuracy']

training_accuracy7 = history7.history['sparse_categorical_accuracy']
validation_accuracy7 = history7.history['val_sparse_categorical_accuracy']

training_accuracy8 = history8.history['sparse_categorical_accuracy']
validation_accuracy8 = history8.history['val_sparse_categorical_accuracy']

epochs_range=range(100)

plt.figure(figsize=(8, 8))
plt.subplot(1, 2, 1)
plt.plot(epochs_range, training_accuracy0, label='Train Acc for Baseline')
plt.plot(epochs_range, training_accuracy1, label='Train Acc for Model1')
plt.plot(epochs_range, training_accuracy2, label='Train Acc for Model2')
plt.plot(epochs_range, training_accuracy3, label='Train Acc for Model3')
plt.plot(epochs_range, training_accuracy4, label='Train Acc for Model4')
plt.plot(epochs_range, training_accuracy5, label='Train Acc for Model5')
plt.plot(epochs_range, training_accuracy6, label='Train Acc for Model6')
plt.plot(epochs_range, training_accuracy7, label='Train Acc for Model7')
plt.plot(epochs_range, training_accuracy8, label='Train Acc for Model8')
plt.legend(loc='lower right')
plt.title('Training Accuracy For All Models')

plt.subplot(1, 2, 2)
plt.plot(epochs_range, validation_accuracy0, label='Val Acc for Baseline')
plt.plot(epochs_range, validation_accuracy1, label='Val Acc for Model1')
plt.plot(epochs_range, validation_accuracy2, label='Val Acc for Model2')
plt.plot(epochs_range, validation_accuracy3, label='Val Acc for Model3')
plt.plot(epochs_range, validation_accuracy4, label='Val Acc for Model4')
plt.plot(epochs_range, validation_accuracy5, label='Val Acc for Model5')
plt.plot(epochs_range, validation_accuracy6, label='Val Acc for Model6')
plt.plot(epochs_range, validation_accuracy7, label='Val Acc for Model7')
plt.plot(epochs_range, validation_accuracy8, label='Val Acc for Model8')
plt.legend(loc='lower right')
plt.title('Validation Accuracy For All Models')
plt.show()
```


Training Accuracy For All Models
Train Acc for Baseline
Train Acc for Model1
Train Acc for Model2
Train Acc for Model3
Train Acc for Model4
Train Acc for Model5
Train Acc for Model6
Train Acc for Model7
Train Acc for Model8


Validation Accuracy For All Models
Val Acc for Baseline
Val Acc for Model1
Val Acc for Model2
Val Acc for Model3
Val Acc for Model4
Val Acc for Model5
Val Acc for Model6
Val Acc for Model7
Val Acc for Model8