RNN for Human Activity Recognition - 2D Pose Input

This experiment is the classification of human activities using a 2D pose time series dataset and an LSTM RNN. The idea is to prove the concept that using a series of 2D poses, rather than 3D poses or a raw 2D images, can produce an accurate estimation of the behaviour of a person or animal. This is a step towards creating a method of classifying an animal's current behaviour state and predicting it's likely next state, allowing for better interaction with an autonomous mobile robot.

Objectives

The aims of this experiment are:

- To determine if 2D pose has comparable accuracy to 3D pose for use in activity recognition. This would allow the use of RGB only cameras for human and animal pose estimation, as opposed to RGBD or a large motion capture dataset.
- To determine if 2D pose has comparable accuracy to using raw RGB images for use in activity recognition. This is based on the idea that limiting the input feature vector can help to deal with a limited dataset, as is likely to occur in animal activity recognition, by allowing for a smaller model to be used (citation required).
- To verify the concept for use in future works involving behaviour prediction from motion in 2D images.

The network used in this experiment is based on that of Guillaume Chevalier, 'LSTMs for Human Activity Recognition, 2016' https://github.com/guillaume-chevalier/LSTM-Human-Activity-Recognition (https://github.com/guillaume-chevalier/LSTM-Human-Activity-Recognition), available under the MIT License. Notable changes that have been made (other than accounting for dataset sizes) are:

- Adapting for use with a large dataset ordered by class, using random sampling without replacement for mini-batch.
  This allows for use of smaller batch sizes when using a dataset ordered by class. "It has been observed in practice that when using a larger batch there is a significant degradation in the quality of the model, as measured by its ability to generalize"
  N.S Keskar, D. Mudigere, et al, 'On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima', ICLR 2017 https://arxiv.org/abs/1609.04836 (https://arxiv.org/abs/1609.04836)
- Exponentially decaying learning rate implemented

Dataset overview

The dataset consists of pose estimations, made using the software OpenPose (https://github.com/CMU-Perceptual-Computing-Lab/openpose's (https://github.com/CMU-Perceptual-Computing-Lab/openpose's)) on a subset of the Berkeley Multimodal Human Action Database (MHAD) dataset http://tele-immersion.citris-uc.org/berkeley_mhad (http://tele-immersion.citris-uc.org/berkeley_mhad).

This dataset is comprised of 12 subjects doing the following 6 actions for 5 repetitions, filmed from 4 angles, repeated 5 times each.

- JUMPING,
- JUMPING_JACKS,
- BOXING,
- WAVING_2HANDS,
- WAVING_1HAND,
- CLAPPING_HANDS.

In total, there are 1438 videos (2 were missing) made up of 211200 individual frames.

The below image is an example of the 4 camera views during the 'boxing' action for subject 1

 alt text

The input for the LSTM is the 2D position of 18 joints across a timeseries of frames numbering n_steps (window-width), with an associated class label for the frame series.
A single frame's input (where j refers to a joint) is stored as:

[j0_x, j0_y, j1_x, j1_y , j2_x, j2_y, j3_x, j3_y, j4_x, j4_y, j5_x, j5_y, j6_x, j6_y, j7_x, j7_y, j8_x, j8_y, j9_x, j9_y, j10_x, j10_y, j11_x, j11_y, j12_x, j12_y, j13_x, j13_y, j14_x, j14_y, j15_x, j15_y, j16_x, j16_y, j17_x, j17_y]

For the following experiment, very little preprocessing has been done to the dataset.
The following steps were taken:

1. openpose run on individual frames, for each subject, action and view, outputting JSON of 18 joint x and y position keypoints and accuracies per frame
2. JSONs converted into txt format, keeping only x and y positions of each frame, action being performed during frame, and order of frames. This is used to create a database of associated activity class number and corresponding series of joint 2D positions
3. No further prepossessing was performed.

In some cases, multiple people were detected in each frame, in which only the first detection was used.

The data has not been normalised with regards to subject position in the frame, motion across frame (if any), size of the subject, speed of action etc. It is essentially the raw 2D position of each joint viewed from a stationary camera.
In many cases, individual joints were not located and a position of [0.0,0.0] was given for that joint

A summary of the dataset used for input is:

- 211200 individual images
- n_steps = 32 frames (~=1.5s at 22Hz)
- Images with noisy pose detection (detection of >=2 people) = 5132
- Training_split = 0.8
- Overlap = 0.8125 (26 / 32) ie 26 frame overlap
    - Length X_train = 22625 * 32 frames
    - Length X_test = 5751 * 32 frames

Note that their is no overlap between test and train sets, which were seperated by activity repetition entirely, before creating the 26 of 32 frame overlap.

Training and Results below:

Training took approximately 4 mins running on a single GTX1080Ti, and was run for 22,000,000ish iterations with a batch size of 5000 (600 epochs)

In [9]:

```
# useful tricks
# to suppress output of each cell; https://stackoverflow.com/questions/23692950/how-do-
you-suppress-output-in-ipython-notebook

import numpy as np
import tensorflow as tf
import matplotlib # not in the environment previously
import matplotlib.pyplot as plt
#import tensorflow as tf  # Version 1.0.0 (some previous versions are used in past comm
its)
#import tensorflow.compat.v1 as tf
#tf.disable_v2_behavior()
from sklearn import metrics# not in the environment previously
import random
from random import randint
import time
import os

import warnings
# reference - https://stackoverflow.com/questions/9031783/hide-all-warnings-in-ipython/
9031848
# warnings.filterwarnings('ignore') # suppress;
warnings.filterwarnings(action='once') # display the warnings once;

# check for gpu access?
tf.Session(config=tf.ConfigProto(log_device_placement=True))
print(tf.VERSION)
```

```
1.13.1
```

Preparing dataset:

In [3]:

```
# Useful Constants

# Output classes to learn how to classify
LABELS = [
    "JUMPING",
    "JUMPING_JACKS",
    "BOXING",
    "WAVING_2HANDS",
    "WAVING_1HAND",
    "CLAPPING_HANDS"

]
DATASET_PATH = "dataset/"

X_train_path = DATASET_PATH + "X_train.txt"
X_test_path = DATASET_PATH + "X_test.txt"

y_train_path = DATASET_PATH + "Y_train.txt"
y_test_path = DATASET_PATH + "Y_test.txt"

n_steps = 32 # 32 timesteps per series
```

In [4]:

```
# Load the networks inputs

def load_X(X_path):
    file = open(X_path, 'r')
    X_ = np.array(
        [elem for elem in [
            row.split(',') for row in file
        ]],
        dtype=np.float32
    )
    file.close()
    blocks = int(len(X_) / n_steps)

    X_ = np.array(np.split(X_,blocks))

    return X_

# Load the networks outputs

def load_y(y_path):
    file = open(y_path, 'r')
    y_ = np.array(
        [elem for elem in [
            row.replace('  ', ' ').strip().split(' ') for row in file
        ]],
        dtype=np.int32
    )
    file.close()

    # for 0-based indexing
    return y_ - 1

X_train = load_X(X_train_path)
X_test = load_X(X_test_path)
#print X_test

y_train = load_y(y_train_path)
y_test = load_y(y_test_path)
# proof that it actually works for the skeptical: replace labelled classes with random
 classes to train on
#for i in range(len(y_train)):
#    y_train[i] = randint(0, 5)
```

Set Parameters:

In [5]:

```
# Input Data

training_data_count = len(X_train)  # 4519 training series (with 50% overlap between ea
ch serie)
test_data_count = len(X_test)  # 1197 test series
n_input = len(X_train[0][0])  # num input parameters per timestep

n_hidden = 34 # Hidden layer num of features
n_classes = 6

#updated for learning-rate decay
# calculated as: decayed_learning_rate = learning_rate * decay_rate ^ (global_step / de
cay_steps)
decaying_learning_rate = True
learning_rate = 0.0025 #used if decaying_learning_rate set to False
init_learning_rate = 0.005
decay_rate = 0.96 #the base of the exponential in the decay
decay_steps = 100000 #used in decay every 60000 steps with a base of 0.96

global_step = tf.Variable(0, trainable=False)
lambda_loss_amount = 0.0015

training_iters = training_data_count *300  # Loop 300 times on the dataset, ie 300 epoc
hs
batch_size = 512
display_iter = batch_size*8  # To show test set accuracy during training

print("(X shape, y shape, every X's mean, every X's standard deviation)")
print(X_train.shape, y_test.shape, np.mean(X_test), np.std(X_test))
print("\nThe dataset has not been preprocessed, is not normalised etc\n")
print('hello\n')
```

```
WARNING:tensorflow:From C:\ProgramData\Anaconda3\envs\tf_gpu\lib\site-pack
ages\tensorflow\python\framework\op_def_library.py:263: colocate_with (fro
m tensorflow.python.framework.ops) is deprecated and will be removed in a
future version.
Instructions for updating:
Colocations handled automatically by placer.
(X shape, y shape, every X's mean, every X's standard deviation)
(22625, 32, 36) (5751, 1) 251.01117 126.12204

The dataset has not been preprocessed, is not normalised etc

hello
```

Utility functions for training:

In [6]:

```
def LSTM_RNN(_X, _weights, _biases):
    # model architecture based on "guillaume-chevalier" and "aymericdamien" under the M
IT license.

    _X = tf.transpose(_X, [1, 0, 2])  # permute n_steps and batch_size
    _X = tf.reshape(_X, [-1, n_input])
    # Rectifies Linear Unit activation function used
    _X = tf.nn.relu(tf.matmul(_X, _weights['hidden']) + _biases['hidden'])
    # Split data because rnn cell needs a list of inputs for the RNN inner loop
    _X = tf.split(_X, n_steps, 0)

    # Define two stacked LSTM cells (two recurrent layers deep) with tensorflow
    lstm_cell_1 = tf.contrib.rnn.BasicLSTMCell(n_hidden, forget_bias=1.0, state_is_tupl
e=True)
    lstm_cell_2 = tf.contrib.rnn.BasicLSTMCell(n_hidden, forget_bias=1.0, state_is_tupl
e=True)
    lstm_cells = tf.contrib.rnn.MultiRNNCell([lstm_cell_1, lstm_cell_2], state_is_tuple
=True)
    outputs, states = tf.contrib.rnn.static_rnn(lstm_cells, _X, dtype=tf.float32)

    # A single output is produced, in style of "many to one" classifier, refer to htt
p://karpathy.github.io/2015/05/21/rnn-effectiveness/ for details
    lstm_last_output = outputs[-1]

    # Linear activation
    return tf.matmul(lstm_last_output, _weights['out']) + _biases['out']


def extract_batch_size(_train, _labels, _unsampled, batch_size):
    # Fetch a "batch_size" amount of data and labels from "(X|y)_train" data.
    # Elements of each batch are chosen randomly, without replacement, from X_train wit
h corresponding label from Y_train
    # unsampled_indices keeps track of sampled data ensuring non-replacement. Resets wh
en remaining datapoints < batch_size

    shape = list(_train.shape)
    shape[0] = batch_size
    batch_s = np.empty(shape)
    batch_labels = np.empty((batch_size,1))
    for i in range(batch_size):
        # Loop index
        # index = random sample from _unsampled (indices)
        index = random.choice(_unsampled)
        batch_s[i] = _train[index]
        batch_labels[i] = _labels[index]
        # yick-modified;
        # reference - https://stackoverflow.com/questions/28150965/why-range0-10-remove
1-does-not-work
        #_unsampled.remove(index) # note: '_unsampled' is of class: range; see the refe
rence;
        _unsampled = [i for i in _unsampled if i != index]

    return batch_s, batch_labels, _unsampled


def one_hot(y_):
    # One hot encoding of the network outputs
    # e.g.: [[5], [0], [3]] --> [[0, 0, 0, 0, 0, 1], [1, 0, 0, 0, 0, 0], [0, 0, 0, 1,
 0, 0]]
```

```
    y_ = y_.reshape(len(y_))
    n_values = int(np.max(y_)) + 1
    return np.eye(n_values)[np.array(y_, dtype=np.int32)]  # Returns FLOATS
```

Build the network:

In [7]:

```
# Graph input/output
# warning - the attribute: "placeholder" is deprecated for tensorflow versions > 2
# https://better-coding.com/solved-attributeerror-module-tensorflow-has-no-attribute-placeholder/

x = tf.placeholder(tf.float32, [None, n_steps, n_input])
y = tf.placeholder(tf.float32, [None, n_classes])

# Graph weights
weights = {
    'hidden': tf.Variable(tf.random_normal([n_input, n_hidden])), # Hidden layer weights
    'out': tf.Variable(tf.random_normal([n_hidden, n_classes], mean=1.0))
}
biases = {
    'hidden': tf.Variable(tf.random_normal([n_hidden])),
    'out': tf.Variable(tf.random_normal([n_classes]))
}

pred = LSTM_RNN(x, weights, biases)

# Loss, optimizer and evaluation
l2 = lambda_loss_amount * sum(
    tf.nn.l2_loss(tf_var) for tf_var in tf.trainable_variables()
) # L2 loss prevents this overkill neural network to overfit the data
cost = tf.reduce_mean(tf.nn.softmax_cross_entropy_with_logits(labels=y, logits=pred)) + l2 # Softmax loss
if decaying_learning_rate:
    learning_rate = tf.train.exponential_decay(init_learning_rate, global_step*batch_size, decay_steps, decay_rate, staircase=True)


#decayed_learning_rate = learning_rate * decay_rate ^ (global_step / decay_steps) #exponentially decayed learning rate
optimizer = tf.train.AdamOptimizer(learning_rate=learning_rate).minimize(cost,global_step=global_step) # Adam Optimizer

correct_pred = tf.equal(tf.argmax(pred,1), tf.argmax(y,1))
accuracy = tf.reduce_mean(tf.cast(correct_pred, tf.float32))
```

```
WARNING: The TensorFlow contrib module will not be included in TensorFlow
2.0.
For more information, please see:
  * https://github.com/tensorflow/community/blob/master/rfcs/20180907-cont
rib-sunset.md
  * https://github.com/tensorflow/addons
If you depend on functionality not listed there, please file an issue.

WARNING:tensorflow:From <ipython-input-6-27f17d3c2292>:12: BasicLSTMCell._
_init__ (from tensorflow.python.ops.rnn_cell_impl) is deprecated and will
be removed in a future version.
Instructions for updating:
This class is equivalent as tf.keras.layers.LSTMCell, and will be replaced
by that in Tensorflow 2.0.
WARNING:tensorflow:From <ipython-input-6-27f17d3c2292>:14: MultiRNNCell.__
init__ (from tensorflow.python.ops.rnn_cell_impl) is deprecated and will b
e removed in a future version.
Instructions for updating:
This class is equivalent as tf.keras.layers.StackedRNNCells, and will be r
eplaced by that in Tensorflow 2.0.
WARNING:tensorflow:From <ipython-input-6-27f17d3c2292>:15: static_rnn (fro
m tensorflow.python.ops.rnn) is deprecated and will be removed in a future
version.
Instructions for updating:
Please use `keras.layers.RNN(cell, unroll=True)`, which is equivalent to t
his API
WARNING:tensorflow:From <ipython-input-7-f3a6715feeef>:24: softmax_cross_e
ntropy_with_logits (from tensorflow.python.ops.nn_ops) is deprecated and w
ill be removed in a future version.
Instructions for updating:

Future major versions of TensorFlow will allow gradients to flow
into the labels input on backprop by default.

See `tf.nn.softmax_cross_entropy_with_logits_v2`.
```

Train the network:

In [8]:

```
test_losses = []
test_accuracies = []
train_losses = []
train_accuracies = []
sess = tf.InteractiveSession(config=tf.ConfigProto(log_device_placement=True))
init = tf.global_variables_initializer()
sess.run(init)

# Perform Training steps with "batch_size" amount of data at each loop.
# Elements of each batch are chosen randomly, without replacement, from X_train,
# restarting when remaining datapoints < batch_size
step = 1
time_start = time.time()
unsampled_indices = range(0,len(X_train))

while step * batch_size <= training_iters:
    #print (sess.run(learning_rate)) #decaying learning rate
    #print (sess.run(global_step)) # global number of iterations
    if len(unsampled_indices) < batch_size:
        unsampled_indices = range(0,len(X_train))
    batch_xs, raw_labels, unsampled_indicies = extract_batch_size(X_train, y_train, uns
ampled_indices, batch_size)
    batch_ys = one_hot(raw_labels)
    # check that encoded output is same length as num_classes, if not, pad it
    if len(batch_ys[0]) < n_classes:
        temp_ys = np.zeros((batch_size, n_classes))
        temp_ys[:batch_ys.shape[0],:batch_ys.shape[1]] = batch_ys
        batch_ys = temp_ys



    # Fit training using batch data
    _, loss, acc = sess.run(
        [optimizer, cost, accuracy],
        feed_dict={
            x: batch_xs,
            y: batch_ys
        }
    )
    train_losses.append(loss)
    train_accuracies.append(acc)

    # Evaluate network only at some steps for faster training:
    if (step*batch_size % display_iter == 0) or (step == 1) or (step * batch_size > tra
ining_iters):

        # To not spam console, show training accuracy/loss in this "if"
        print("Iter #" + str(step*batch_size) + \
              ":  Learning rate = " + "{:.6f}".format(sess.run(learning_rate)) + \
              ":   Batch Loss = " + "{:.6f}".format(loss) + \
              ", Accuracy = {}".format(acc))

        # Evaluation on the test set (no learning made here - just evaluation for diagn
osis)
        loss, acc = sess.run(
            [cost, accuracy],
            feed_dict={
                x: X_test,
                y: one_hot(y_test)
```

```
            }
        )
        test_losses.append(loss)
        test_accuracies.append(acc)
        print("PERFORMANCE ON TEST SET:             " + \
              "Batch Loss = {}".format(loss) + \
              ", Accuracy = {}".format(acc))

    step += 1

print("Optimization Finished!")

# Accuracy for test data

one_hot_predictions, accuracy, final_loss = sess.run(
    [pred, accuracy, cost],
    feed_dict={
        x: X_test,
        y: one_hot(y_test)
    }
)

test_losses.append(final_loss)
test_accuracies.append(accuracy)

print("FINAL RESULT: " + \
      "Batch Loss = {}".format(final_loss) + \
      ", Accuracy = {}".format(accuracy))
time_stop = time.time()
print("TOTAL TIME:  {}".format(time_stop - time_start))
```

```
Iter #512:  Learning rate = 0.005000:   Batch Loss = 3.526922, Accuracy =
0.19921875
PERFORMANCE ON TEST SET:             Batch Loss = 3.3879947662353516, Accu
racy = 0.2197878658771515
Iter #4096:  Learning rate = 0.005000:   Batch Loss = 2.989375, Accuracy =
0.306640625
PERFORMANCE ON TEST SET:             Batch Loss = 2.887584686279297, Accur
acy = 0.34307077527046204
Iter #8192:  Learning rate = 0.005000:   Batch Loss = 2.723924, Accuracy =
0.404296875
PERFORMANCE ON TEST SET:             Batch Loss = 2.683004379272461, Accur
acy = 0.4336636960506439
Iter #12288:  Learning rate = 0.005000:   Batch Loss = 2.510427, Accuracy
 = 0.50390625
PERFORMANCE ON TEST SET:             Batch Loss = 2.4980244636535645, Accu
racy = 0.5416449308395386
Iter #16384:  Learning rate = 0.005000:   Batch Loss = 2.549774, Accuracy
 = 0.416015625
PERFORMANCE ON TEST SET:             Batch Loss = 2.5183496475219727, Accu
racy = 0.45852896571159363
Iter #20480:  Learning rate = 0.005000:   Batch Loss = 2.460151, Accuracy
 = 0.41796875
PERFORMANCE ON TEST SET:             Batch Loss = 2.3421549797058105, Accu
racy = 0.5367761850357056
Iter #24576:  Learning rate = 0.005000:   Batch Loss = 2.489107, Accuracy
 = 0.43359375
PERFORMANCE ON TEST SET:             Batch Loss = 2.3953800201416016, Accu
racy = 0.49052339792251587
Iter #28672:  Learning rate = 0.005000:   Batch Loss = 2.322978, Accuracy
 = 0.501953125
PERFORMANCE ON TEST SET:             Batch Loss = 2.2568039894104004, Accu
racy = 0.5263432264328003
Iter #32768:  Learning rate = 0.005000:   Batch Loss = 2.097795, Accuracy
 = 0.62890625
PERFORMANCE ON TEST SET:             Batch Loss = 2.1531424522399902, Accu
racy = 0.5722482800483704
Iter #36864:  Learning rate = 0.005000:   Batch Loss = 2.243058, Accuracy
 = 0.54296875
PERFORMANCE ON TEST SET:             Batch Loss = 2.0515410900115967, Accu
racy = 0.6167622804641724
Iter #40960:  Learning rate = 0.005000:   Batch Loss = 2.080172, Accuracy
 = 0.578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.9987791776657104, Accu
racy = 0.6131107807159424
Iter #45056:  Learning rate = 0.005000:   Batch Loss = 2.038051, Accuracy
 = 0.564453125
PERFORMANCE ON TEST SET:             Batch Loss = 1.9346234798431396, Accu
racy = 0.6298035383224487
Iter #49152:  Learning rate = 0.005000:   Batch Loss = 1.928234, Accuracy
 = 0.61328125
PERFORMANCE ON TEST SET:             Batch Loss = 1.9215298891067505, Accu
racy = 0.656581461429596
Iter #53248:  Learning rate = 0.005000:   Batch Loss = 1.875210, Accuracy
 = 0.697265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.859198808670044, Accur
acy = 0.6784906983375549
Iter #57344:  Learning rate = 0.005000:   Batch Loss = 1.904188, Accuracy
 = 0.626953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.876065969467163, Accur
acy = 0.6548426151275635
Iter #61440:  Learning rate = 0.005000:   Batch Loss = 1.718810, Accuracy
```

```
 = 0.70703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.8030116558074951, Accu
racy = 0.6701443195343018
Iter #65536:  Learning rate = 0.005000:   Batch Loss = 1.770049, Accuracy
 = 0.67578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.799513339996338, Accur
acy = 0.6857937574386597
Iter #69632:  Learning rate = 0.005000:   Batch Loss = 1.811988, Accuracy
 = 0.669921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.7900855541229248, Accu
racy = 0.687706470489502
Iter #73728:  Learning rate = 0.005000:   Batch Loss = 1.737167, Accuracy
 = 0.724609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.7375303506851196, Accu
racy = 0.6991827487945557
Iter #77824:  Learning rate = 0.005000:   Batch Loss = 1.726960, Accuracy
 = 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.7098560333251953, Accu
racy = 0.7160493731498718
Iter #81920:  Learning rate = 0.005000:   Batch Loss = 1.677000, Accuracy
 = 0.6953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.756859540939331, Accur
acy = 0.6746652722358704
Iter #86016:  Learning rate = 0.005000:   Batch Loss = 1.693488, Accuracy
 = 0.712890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.6646754741668701, Accu
racy = 0.7200486660003662
Iter #90112:  Learning rate = 0.005000:   Batch Loss = 1.634446, Accuracy
 = 0.6796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.6393654346466064, Accu
racy = 0.7101373672485352
Iter #94208:  Learning rate = 0.005000:   Batch Loss = 1.731565, Accuracy
 = 0.66015625
PERFORMANCE ON TEST SET:             Batch Loss = 1.7006075382232666, Accu
racy = 0.6715353727340698
Iter #98304:  Learning rate = 0.005000:   Batch Loss = 1.557584, Accuracy
 = 0.728515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.7053958177566528, Accu
racy = 0.6899669766426086
Iter #102400:  Learning rate = 0.004800:   Batch Loss = 1.607659, Accuracy
= 0.71484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.5777344703674316, Accu
racy = 0.7089201807975769
Iter #106496:  Learning rate = 0.004800:   Batch Loss = 1.560118, Accuracy
= 0.728515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.5135005712509155, Accu
racy = 0.7440445423126221
Iter #110592:  Learning rate = 0.004800:   Batch Loss = 1.585845, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.521201729774475, Accur
acy = 0.729264497756958
Iter #114688:  Learning rate = 0.004800:   Batch Loss = 1.507756, Accuracy
= 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.5903515815734863, Accu
racy = 0.7125717401504517
Iter #118784:  Learning rate = 0.004800:   Batch Loss = 1.470733, Accuracy
= 0.736328125
PERFORMANCE ON TEST SET:             Batch Loss = 1.523011565208435, Accur
acy = 0.7348287105560303
Iter #122880:  Learning rate = 0.004800:   Batch Loss = 1.562328, Accuracy
= 0.705078125
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 1.5471266508102417, Accu
racy = 0.7040514945983887
Iter #126976:  Learning rate = 0.004800:   Batch Loss = 1.523246, Accuracy
= 0.720703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.6301023960113525, Accu
racy = 0.6491044759750366
Iter #131072:  Learning rate = 0.004800:   Batch Loss = 1.540357, Accuracy
= 0.71484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.522336483001709, Accur
acy = 0.7304816842079163
Iter #135168:  Learning rate = 0.004800:   Batch Loss = 1.446889, Accuracy
= 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.441519856452942, Accur
acy = 0.7381324768066406
Iter #139264:  Learning rate = 0.004800:   Batch Loss = 1.424098, Accuracy
= 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.5513629913330078, Accu
racy = 0.6974439024925232
Iter #143360:  Learning rate = 0.004800:   Batch Loss = 1.500430, Accuracy
= 0.740234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.5047171115875244, Accu
racy = 0.7115284204483032
Iter #147456:  Learning rate = 0.004800:   Batch Loss = 1.575858, Accuracy
= 0.69140625
PERFORMANCE ON TEST SET:             Batch Loss = 1.583299160003662, Accur
acy = 0.6960528492927551
Iter #151552:  Learning rate = 0.004800:   Batch Loss = 1.429147, Accuracy
= 0.744140625
PERFORMANCE ON TEST SET:             Batch Loss = 1.4860950708389282, Accu
racy = 0.7221353054046631
Iter #155648:  Learning rate = 0.004800:   Batch Loss = 1.395472, Accuracy
= 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 1.585444450378418, Accur
acy = 0.6967483758926392
Iter #159744:  Learning rate = 0.004800:   Batch Loss = 1.472334, Accuracy
= 0.70703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.4661788940429688, Accu
racy = 0.7209181189537048
Iter #163840:  Learning rate = 0.004800:   Batch Loss = 1.390318, Accuracy
= 0.748046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.4382503032684326, Accu
racy = 0.7541297078132629
Iter #167936:  Learning rate = 0.004800:   Batch Loss = 1.347453, Accuracy
= 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.4505935907363892, Accu
racy = 0.7350026369094849
Iter #172032:  Learning rate = 0.004800:   Batch Loss = 1.433962, Accuracy
= 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.3975821733474731, Accu
racy = 0.7607372403144836
Iter #176128:  Learning rate = 0.004800:   Batch Loss = 1.327404, Accuracy
= 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3577251434326172, Accu
racy = 0.7753434181213379
Iter #180224:  Learning rate = 0.004800:   Batch Loss = 1.293792, Accuracy
= 0.79296875
PERFORMANCE ON TEST SET:             Batch Loss = 1.3613824844360352, Accu
racy = 0.7537819743156433
Iter #184320:  Learning rate = 0.004800:   Batch Loss = 1.360129, Accuracy
= 0.740234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3577256202697754, Accu
```

```
racy = 0.7576074004173279
Iter #188416:  Learning rate = 0.004800:   Batch Loss = 1.267964, Accuracy
= 0.806640625
PERFORMANCE ON TEST SET:             Batch Loss = 1.470594882965088, Accur
acy = 0.729960024356842
Iter #192512:  Learning rate = 0.004800:   Batch Loss = 1.310650, Accuracy
= 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3367326259613037, Accu
racy = 0.7657798528671265
Iter #196608:  Learning rate = 0.004800:   Batch Loss = 1.281442, Accuracy
= 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3170535564422607, Accu
racy = 0.7703008055686951
Iter #200704:  Learning rate = 0.004608:   Batch Loss = 1.330364, Accuracy
= 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3164691925048828, Accu
racy = 0.7715179920196533
Iter #204800:  Learning rate = 0.004608:   Batch Loss = 1.259244, Accuracy
= 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 3.7963056564331055, Accu
racy = 0.4708746373653412
Iter #208896:  Learning rate = 0.004608:   Batch Loss = 2.019292, Accuracy
= 0.4453125
PERFORMANCE ON TEST SET:             Batch Loss = 2.0307509899139404, Accu
racy = 0.46078944206237793
Iter #212992:  Learning rate = 0.004608:   Batch Loss = 1.955724, Accuracy
= 0.48046875
PERFORMANCE ON TEST SET:             Batch Loss = 2.1418728828430176, Accu
racy = 0.47174403071403503
Iter #217088:  Learning rate = 0.004608:   Batch Loss = 1.904222, Accuracy
= 0.50390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.8586593866348267, Accu
racy = 0.543383777141571
Iter #221184:  Learning rate = 0.004608:   Batch Loss = 1.920824, Accuracy
= 0.51953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.7989052534103394, Accu
racy = 0.5929403305053711
Iter #225280:  Learning rate = 0.004608:   Batch Loss = 1.734341, Accuracy
= 0.587890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.796802282333374, Accur
acy = 0.559728741645813
Iter #229376:  Learning rate = 0.004608:   Batch Loss = 1.755063, Accuracy
= 0.572265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.7490395307540894, Accu
racy = 0.5927664637565613
Iter #233472:  Learning rate = 0.004608:   Batch Loss = 3.869408, Accuracy
= 0.29296875
PERFORMANCE ON TEST SET:             Batch Loss = 3.153437852859497, Accur
acy = 0.32116153836250305
Iter #237568:  Learning rate = 0.004608:   Batch Loss = 2.379040, Accuracy
= 0.2421875
PERFORMANCE ON TEST SET:             Batch Loss = 2.321929454803467, Accur
acy = 0.24726134538650513
Iter #241664:  Learning rate = 0.004608:   Batch Loss = 2.282774, Accuracy
= 0.39453125
PERFORMANCE ON TEST SET:             Batch Loss = 2.2064456939697266, Accu
racy = 0.44983482360839844
Iter #245760:  Learning rate = 0.004608:   Batch Loss = 2.255287, Accuracy
= 0.35546875
PERFORMANCE ON TEST SET:             Batch Loss = 2.136434316635132, Accur
acy = 0.46826639771461487
```

```
Iter #249856:  Learning rate = 0.004608:   Batch Loss = 2.155848, Accuracy
= 0.40234375
PERFORMANCE ON TEST SET:             Batch Loss = 2.0976715087890625, Accu
racy = 0.4315771162509918
Iter #253952:  Learning rate = 0.004608:   Batch Loss = 2.139487, Accuracy
= 0.41796875
PERFORMANCE ON TEST SET:             Batch Loss = 2.003925085067749, Accur
acy = 0.48495912551879883
Iter #258048:  Learning rate = 0.004608:   Batch Loss = 2.119453, Accuracy
= 0.419921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.9923155307769775, Accu
racy = 0.4891323149204254
Iter #262144:  Learning rate = 0.004608:   Batch Loss = 2.173018, Accuracy
= 0.416015625
PERFORMANCE ON TEST SET:             Batch Loss = 2.0595149993896484, Accu
racy = 0.45244306325912476
Iter #266240:  Learning rate = 0.004608:   Batch Loss = 2.131440, Accuracy
= 0.447265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.9113222360610962, Accu
racy = 0.5425143241882324
Iter #270336:  Learning rate = 0.004608:   Batch Loss = 1.990546, Accuracy
= 0.501953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.829270839691162, Accur
acy = 0.5656407475471497
Iter #274432:  Learning rate = 0.004608:   Batch Loss = 1.888039, Accuracy
= 0.494140625
PERFORMANCE ON TEST SET:             Batch Loss = 1.883558988571167, Accur
acy = 0.5647713541984558
Iter #278528:  Learning rate = 0.004608:   Batch Loss = 1.793662, Accuracy
= 0.6171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.7735941410064697, Accu
racy = 0.5818118453025818
Iter #282624:  Learning rate = 0.004608:   Batch Loss = 1.772268, Accuracy
= 0.59765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.7108302116394043, Accu
racy = 0.6077203750610352
Iter #286720:  Learning rate = 0.004608:   Batch Loss = 2.221627, Accuracy
= 0.466796875
PERFORMANCE ON TEST SET:             Batch Loss = 2.105194091796875, Accur
acy = 0.4061902165412903
Iter #290816:  Learning rate = 0.004608:   Batch Loss = 1.851532, Accuracy
= 0.51953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.7904338836669922, Accu
racy = 0.5663362741470337
Iter #294912:  Learning rate = 0.004608:   Batch Loss = 1.727595, Accuracy
= 0.61328125
PERFORMANCE ON TEST SET:             Batch Loss = 1.684667944908142, Accur
acy = 0.6271952986717224
Iter #299008:  Learning rate = 0.004608:   Batch Loss = 1.706117, Accuracy
= 0.587890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.6852096319198608, Accu
racy = 0.6318901181221008
Iter #303104:  Learning rate = 0.004424:   Batch Loss = 1.621914, Accuracy
= 0.68359375
PERFORMANCE ON TEST SET:             Batch Loss = 1.6107242107391357, Accu
racy = 0.6567553281784058
Iter #307200:  Learning rate = 0.004424:   Batch Loss = 1.713464, Accuracy
= 0.591796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.6684670448303223, Accu
racy = 0.6504955887794495
Iter #311296:  Learning rate = 0.004424:   Batch Loss = 1.881243, Accuracy
```

```
= 0.5859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.8611650466918945, Accu
racy = 0.5772908926010132
Iter #315392:  Learning rate = 0.004424:   Batch Loss = 1.693546, Accuracy
= 0.658203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.7866342067718506, Accu
racy = 0.6016345024108887
Iter #319488:  Learning rate = 0.004424:   Batch Loss = 1.677834, Accuracy
= 0.630859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.7205705642700195, Accu
racy = 0.6285863518714905
Iter #323584:  Learning rate = 0.004424:   Batch Loss = 1.656010, Accuracy
= 0.634765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.9219515323638916, Accu
racy = 0.5574682950973511
Iter #327680:  Learning rate = 0.004424:   Batch Loss = 1.652952, Accuracy
= 0.65234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.823986530303955, Accur
acy = 0.5927664637565613
Iter #331776:  Learning rate = 0.004424:   Batch Loss = 1.947315, Accuracy
= 0.556640625
PERFORMANCE ON TEST SET:             Batch Loss = 1.823476791381836, Accur
acy = 0.5884193778038025
Iter #335872:  Learning rate = 0.004424:   Batch Loss = 1.852860, Accuracy
= 0.560546875
PERFORMANCE ON TEST SET:             Batch Loss = 1.8468177318572998, Accu
racy = 0.5633803009986877
Iter #339968:  Learning rate = 0.004424:   Batch Loss = 1.703680, Accuracy
= 0.6171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.673468828201294, Accur
acy = 0.6320639848709106
Iter #344064:  Learning rate = 0.004424:   Batch Loss = 1.672543, Accuracy
= 0.63671875
PERFORMANCE ON TEST SET:             Batch Loss = 1.5810329914093018, Accu
racy = 0.6765779852867126
Iter #348160:  Learning rate = 0.004424:   Batch Loss = 1.579378, Accuracy
= 0.646484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.6629912853240967, Accu
racy = 0.6270213723182678
Iter #352256:  Learning rate = 0.004424:   Batch Loss = 1.662466, Accuracy
= 0.6484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.5697264671325684, Accu
racy = 0.6830116510391235
Iter #356352:  Learning rate = 0.004424:   Batch Loss = 1.786973, Accuracy
= 0.576171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.839998483657837, Accur
acy = 0.5917232036590576
Iter #360448:  Learning rate = 0.004424:   Batch Loss = 1.630975, Accuracy
= 0.646484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.6236395835876465, Accu
racy = 0.6717092394828796
Iter #364544:  Learning rate = 0.004424:   Batch Loss = 1.553076, Accuracy
= 0.689453125
PERFORMANCE ON TEST SET:             Batch Loss = 1.5695945024490356, Accu
racy = 0.6746652722358704
Iter #368640:  Learning rate = 0.004424:   Batch Loss = 1.603054, Accuracy
= 0.650390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.5194129943847656, Accu
racy = 0.6944879293441772
Iter #372736:  Learning rate = 0.004424:   Batch Loss = 1.530496, Accuracy
= 0.66796875
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 1.5258464813232422, Accu
racy = 0.6685793995857239
Iter #376832:  Learning rate = 0.004424:   Batch Loss = 1.462585, Accuracy
= 0.71875
PERFORMANCE ON TEST SET:             Batch Loss = 1.495157241821289, Accur
acy = 0.6957051157951355
Iter #380928:  Learning rate = 0.004424:   Batch Loss = 1.481490, Accuracy
= 0.70703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.5324163436889648, Accu
racy = 0.6922274231910706
Iter #385024:  Learning rate = 0.004424:   Batch Loss = 1.472206, Accuracy
= 0.703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.4896340370178223, Accu
racy = 0.7061380743980408
Iter #389120:  Learning rate = 0.004424:   Batch Loss = 1.440611, Accuracy
= 0.71484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.5143241882324219, Accu
racy = 0.7009215950965881
Iter #393216:  Learning rate = 0.004424:   Batch Loss = 1.532011, Accuracy
= 0.681640625
PERFORMANCE ON TEST SET:             Batch Loss = 1.5086994171142578, Accu
racy = 0.6917057633399963
Iter #397312:  Learning rate = 0.004424:   Batch Loss = 1.515518, Accuracy
= 0.669921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.4850101470947266, Accu
racy = 0.6668405532836914
Iter #401408:  Learning rate = 0.004247:   Batch Loss = 1.459402, Accuracy
= 0.71484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.5351545810699463, Accu
racy = 0.6802295446395874
Iter #405504:  Learning rate = 0.004247:   Batch Loss = 1.425560, Accuracy
= 0.701171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.5030686855316162, Accu
racy = 0.6887497901916504
Iter #409600:  Learning rate = 0.004247:   Batch Loss = 1.443297, Accuracy
= 0.705078125
PERFORMANCE ON TEST SET:             Batch Loss = 1.4956059455871582, Accu
racy = 0.6995305418968201
Iter #413696:  Learning rate = 0.004247:   Batch Loss = 1.400302, Accuracy
= 0.720703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.4450464248657227, Accu
racy = 0.7111806869506836
Iter #417792:  Learning rate = 0.004247:   Batch Loss = 1.434812, Accuracy
= 0.701171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.464860439300537, Accur
acy = 0.6967483758926392
Iter #421888:  Learning rate = 0.004247:   Batch Loss = 1.415498, Accuracy
= 0.68359375
PERFORMANCE ON TEST SET:             Batch Loss = 1.4545025825500488, Accu
racy = 0.6652756333351135
Iter #425984:  Learning rate = 0.004247:   Batch Loss = 1.582944, Accuracy
= 0.677734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.4459567070007324, Accu
racy = 0.7064858078956604
Iter #430080:  Learning rate = 0.004247:   Batch Loss = 1.443554, Accuracy
= 0.708984375
PERFORMANCE ON TEST SET:             Batch Loss = 1.446462869644165, Accur
acy = 0.7005738019943237
Iter #434176:  Learning rate = 0.004247:   Batch Loss = 1.432167, Accuracy
= 0.708984375
PERFORMANCE ON TEST SET:             Batch Loss = 1.4599529504776, Accurac
```

```
y = 0.7002260684967041
Iter #438272:  Learning rate = 0.004247:   Batch Loss = 1.388052, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.4717826843261719, Accu
racy = 0.6647539734840393
Iter #442368:  Learning rate = 0.004247:   Batch Loss = 1.458447, Accuracy
= 0.6875
PERFORMANCE ON TEST SET:             Batch Loss = 1.419065237045288, Accur
acy = 0.7136150002479553
Iter #446464:  Learning rate = 0.004247:   Batch Loss = 1.434400, Accuracy
= 0.69921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.470787525177002, Accur
acy = 0.6934446096420288
Iter #450560:  Learning rate = 0.004247:   Batch Loss = 1.398946, Accuracy
= 0.71875
PERFORMANCE ON TEST SET:             Batch Loss = 1.4877262115478516, Accu
racy = 0.6911841630935669
Iter #454656:  Learning rate = 0.004247:   Batch Loss = 1.398216, Accuracy
= 0.708984375
PERFORMANCE ON TEST SET:             Batch Loss = 1.4392534494400024, Accu
racy = 0.6715353727340698
Iter #458752:  Learning rate = 0.004247:   Batch Loss = 1.390986, Accuracy
= 0.734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3963892459869385, Accu
racy = 0.7179620862007141
Iter #462848:  Learning rate = 0.004247:   Batch Loss = 1.409698, Accuracy
= 0.703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.4048569202423096, Accu
racy = 0.7120500802993774
Iter #466944:  Learning rate = 0.004247:   Batch Loss = 1.376956, Accuracy
= 0.7109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.383014440536499, Accur
acy = 0.7203964591026306
Iter #471040:  Learning rate = 0.004247:   Batch Loss = 1.359890, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.398223876953125, Accur
acy = 0.7123978734016418
Iter #475136:  Learning rate = 0.004247:   Batch Loss = 1.413579, Accuracy
= 0.70703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.4773385524749756, Accu
racy = 0.6833594441413879
Iter #479232:  Learning rate = 0.004247:   Batch Loss = 1.390810, Accuracy
= 0.716796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.4385709762573242, Accu
racy = 0.6772735118865967
Iter #483328:  Learning rate = 0.004247:   Batch Loss = 1.420403, Accuracy
= 0.693359375
PERFORMANCE ON TEST SET:             Batch Loss = 1.4782801866531372, Accu
racy = 0.6837071776390076
Iter #487424:  Learning rate = 0.004247:   Batch Loss = 1.360901, Accuracy
= 0.732421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.4577922821044922, Accu
racy = 0.693270742893219
Iter #491520:  Learning rate = 0.004247:   Batch Loss = 1.328735, Accuracy
= 0.724609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3893849849700928, Accu
racy = 0.6817944645881653
Iter #495616:  Learning rate = 0.004247:   Batch Loss = 1.375420, Accuracy
= 0.73046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.4012583494186401, Accu
racy = 0.7077029943466187
```

```
Iter #499712:  Learning rate = 0.004247:   Batch Loss = 1.378875, Accuracy
= 0.705078125
PERFORMANCE ON TEST SET:             Batch Loss = 1.3844287395477295, Accu
racy = 0.6826638579368591
Iter #503808:  Learning rate = 0.004077:   Batch Loss = 1.349273, Accuracy
= 0.728515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3844952583312988, Accu
racy = 0.7153538465499878
Iter #507904:  Learning rate = 0.004077:   Batch Loss = 1.435394, Accuracy
= 0.669921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.380551815032959, Accur
acy = 0.6991827487945557
Iter #512000:  Learning rate = 0.004077:   Batch Loss = 1.457085, Accuracy
= 0.703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.464167594909668, Accur
acy = 0.6977916955947876
Iter #516096:  Learning rate = 0.004077:   Batch Loss = 1.423877, Accuracy
= 0.65234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.392505168914795, Accur
acy = 0.6937924027442932
Iter #520192:  Learning rate = 0.004077:   Batch Loss = 1.307522, Accuracy
= 0.74609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.427612066268921, Accur
acy = 0.6979655623435974
Iter #524288:  Learning rate = 0.004077:   Batch Loss = 1.356624, Accuracy
= 0.697265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3945956230163574, Accu
racy = 0.6849243640899658
Iter #528384:  Learning rate = 0.004077:   Batch Loss = 1.320691, Accuracy
= 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 1.3659039735794067, Accu
racy = 0.7170926928520203
Iter #532480:  Learning rate = 0.004077:   Batch Loss = 1.316393, Accuracy
= 0.734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3693712949752808, Accu
racy = 0.7226569056510925
Iter #536576:  Learning rate = 0.004077:   Batch Loss = 1.436801, Accuracy
= 0.6953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.4503333568572998, Accu
racy = 0.6983133554458618
Iter #540672:  Learning rate = 0.004077:   Batch Loss = 1.332697, Accuracy
= 0.73046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.4088377952575684, Accu
racy = 0.6958789825439453
Iter #544768:  Learning rate = 0.004077:   Batch Loss = 1.412349, Accuracy
= 0.708984375
PERFORMANCE ON TEST SET:             Batch Loss = 1.380516529083252, Accur
acy = 0.6868370771408081
Iter #548864:  Learning rate = 0.004077:   Batch Loss = 1.349446, Accuracy
= 0.720703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.3966479301452637, Accu
racy = 0.7080507874488831
Iter #552960:  Learning rate = 0.004077:   Batch Loss = 1.353929, Accuracy
= 0.71875
PERFORMANCE ON TEST SET:             Batch Loss = 1.3982501029968262, Accu
racy = 0.7170926928520203
Iter #557056:  Learning rate = 0.004077:   Batch Loss = 1.296345, Accuracy
= 0.7265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3934152126312256, Accu
racy = 0.7087463140487671
Iter #561152:  Learning rate = 0.004077:   Batch Loss = 1.299976, Accuracy
```

```
= 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.4318385124206543, Accu
racy = 0.7038775682449341
Iter #565248:  Learning rate = 0.004077:   Batch Loss = 1.284779, Accuracy
= 0.7265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.332770586013794, Accur
acy = 0.7304816842079163
Iter #569344:  Learning rate = 0.004077:   Batch Loss = 1.394720, Accuracy
= 0.7265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.5680975914001465, Accu
racy = 0.6704921126365662
Iter #573440:  Learning rate = 0.004077:   Batch Loss = 1.466452, Accuracy
= 0.6875
PERFORMANCE ON TEST SET:             Batch Loss = 1.5358493328094482, Accu
racy = 0.6670144200325012
Iter #577536:  Learning rate = 0.004077:   Batch Loss = 1.450177, Accuracy
= 0.6484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.4972156286239624, Accu
racy = 0.6704921126365662
Iter #581632:  Learning rate = 0.004077:   Batch Loss = 1.253864, Accuracy
= 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.362046480178833, Accur
acy = 0.7249174118041992
Iter #585728:  Learning rate = 0.004077:   Batch Loss = 1.362138, Accuracy
= 0.693359375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3650541305541992, Accu
racy = 0.7181359529495239
Iter #589824:  Learning rate = 0.004077:   Batch Loss = 1.384626, Accuracy
= 0.7109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3897018432617188, Accu
racy = 0.7118762135505676
Iter #593920:  Learning rate = 0.004077:   Batch Loss = 1.358538, Accuracy
= 0.724609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.340620517730713, Accur
acy = 0.7212658524513245
Iter #598016:  Learning rate = 0.004077:   Batch Loss = 1.345024, Accuracy
= 0.703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.3655686378479004, Accu
racy = 0.7191792726516724
Iter #602112:  Learning rate = 0.003914:   Batch Loss = 1.402919, Accuracy
= 0.6875
PERFORMANCE ON TEST SET:             Batch Loss = 1.3637285232543945, Accu
racy = 0.7176143527030945
Iter #606208:  Learning rate = 0.003914:   Batch Loss = 1.312696, Accuracy
= 0.71875
PERFORMANCE ON TEST SET:             Batch Loss = 1.3374688625335693, Accu
racy = 0.6915318965911865
Iter #610304:  Learning rate = 0.003914:   Batch Loss = 1.265388, Accuracy
= 0.736328125
PERFORMANCE ON TEST SET:             Batch Loss = 1.3566665649414062, Accu
racy = 0.7104851603507996
Iter #614400:  Learning rate = 0.003914:   Batch Loss = 1.298002, Accuracy
= 0.728515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.4592437744140625, Accu
racy = 0.6477134227752686
Iter #618496:  Learning rate = 0.003914:   Batch Loss = 1.332345, Accuracy
= 0.70703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.3181686401367188, Accu
racy = 0.7205703258514404
Iter #622592:  Learning rate = 0.003914:   Batch Loss = 1.273949, Accuracy
= 0.7265625
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 1.302191972732544, Accur
acy = 0.7040514945983887
Iter #626688:  Learning rate = 0.003914:   Batch Loss = 1.324779, Accuracy
= 0.724609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3071955442428589, Accu
racy = 0.729960024356842
Iter #630784:  Learning rate = 0.003914:   Batch Loss = 1.224145, Accuracy
= 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3471111059188843, Accu
racy = 0.7172665596008301
Iter #634880:  Learning rate = 0.003914:   Batch Loss = 1.368150, Accuracy
= 0.697265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3166320323944092, Accu
racy = 0.7344809770584106
Iter #638976:  Learning rate = 0.003914:   Batch Loss = 1.323623, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3219902515411377, Accu
racy = 0.6936184763908386
Iter #643072:  Learning rate = 0.003914:   Batch Loss = 1.290074, Accuracy
= 0.73828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.3166389465332031, Accu
racy = 0.7270039916038513
Iter #647168:  Learning rate = 0.003914:   Batch Loss = 1.228023, Accuracy
= 0.744140625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3115233182907104, Accu
racy = 0.7257868051528931
Iter #651264:  Learning rate = 0.003914:   Batch Loss = 1.287390, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2797305583953857, Accu
racy = 0.7369152903556824
Iter #655360:  Learning rate = 0.003914:   Batch Loss = 1.320428, Accuracy
= 0.73046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.408632755279541, Accur
acy = 0.6776213049888611
Iter #659456:  Learning rate = 0.003914:   Batch Loss = 1.254339, Accuracy
= 0.74609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3078505992889404, Accu
racy = 0.7224830389022827
Iter #663552:  Learning rate = 0.003914:   Batch Loss = 1.315511, Accuracy
= 0.70703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.3604899644851685, Accu
racy = 0.7042253613471985
Iter #667648:  Learning rate = 0.003914:   Batch Loss = 1.232205, Accuracy
= 0.73828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.351938247680664, Accur
acy = 0.7122239470481873
Iter #671744:  Learning rate = 0.003914:   Batch Loss = 1.297621, Accuracy
= 0.701171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.332885980606079, Accur
acy = 0.7296122312545776
Iter #675840:  Learning rate = 0.003914:   Batch Loss = 1.362138, Accuracy
= 0.6953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.437676191329956, Accur
acy = 0.6896191835403442
Iter #679936:  Learning rate = 0.003914:   Batch Loss = 1.333456, Accuracy
= 0.70703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.474074363708496, Accur
acy = 0.6770996451377869
Iter #684032:  Learning rate = 0.003914:   Batch Loss = 1.314270, Accuracy
= 0.716796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.3424971103668213, Accu
```

```
racy = 0.7203964591026306
Iter #688128:  Learning rate = 0.003914:   Batch Loss = 1.296818, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2989777326583862, Accu
racy = 0.6937924027442932
Iter #692224:  Learning rate = 0.003914:   Batch Loss = 1.252436, Accuracy
= 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.3069181442260742, Accu
racy = 0.7238740921020508
Iter #696320:  Learning rate = 0.003914:   Batch Loss = 1.262465, Accuracy
= 0.720703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.3739274740219116, Accu
racy = 0.6998782753944397
Iter #700416:  Learning rate = 0.003757:   Batch Loss = 1.248495, Accuracy
= 0.734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3010361194610596, Accu
racy = 0.6983133554458618
Iter #704512:  Learning rate = 0.003757:   Batch Loss = 1.262745, Accuracy
= 0.73828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.3610904216766357, Accu
racy = 0.6950095891952515
Iter #708608:  Learning rate = 0.003757:   Batch Loss = 1.306843, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3312506675720215, Accu
racy = 0.6783168315887451
Iter #712704:  Learning rate = 0.003757:   Batch Loss = 1.220902, Accuracy
= 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.273938536643982, Accur
acy = 0.7330899238586426
Iter #716800:  Learning rate = 0.003757:   Batch Loss = 1.168954, Accuracy
= 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2947503328323364, Accu
racy = 0.7287428379058838
Iter #720896:  Learning rate = 0.003757:   Batch Loss = 1.197111, Accuracy
= 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2778799533843994, Accu
racy = 0.7296122312545776
Iter #724992:  Learning rate = 0.003757:   Batch Loss = 1.303566, Accuracy
= 0.7109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.295098066329956, Accur
acy = 0.701095461845398
Iter #729088:  Learning rate = 0.003757:   Batch Loss = 1.256155, Accuracy
= 0.71875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2959585189819336, Accu
racy = 0.7261345982551575
Iter #733184:  Learning rate = 0.003757:   Batch Loss = 1.206929, Accuracy
= 0.76171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2741310596466064, Accu
racy = 0.7323943376541138
Iter #737280:  Learning rate = 0.003757:   Batch Loss = 1.241812, Accuracy
= 0.74609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2651633024215698, Accu
racy = 0.7155277132987976
Iter #741376:  Learning rate = 0.003757:   Batch Loss = 1.227430, Accuracy
= 0.740234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2359812259674072, Accu
racy = 0.747696042060852
Iter #745472:  Learning rate = 0.003757:   Batch Loss = 1.273331, Accuracy
= 0.744140625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2557384967803955, Accu
racy = 0.7358720302581787
```

```
Iter #749568:  Learning rate = 0.003757:   Batch Loss = 1.232625, Accuracy
= 0.740234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.3666536808013916, Accu
racy = 0.7106590270996094
Iter #753664:  Learning rate = 0.003757:   Batch Loss = 1.320456, Accuracy
= 0.69921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.3334109783172607, Accu
racy = 0.6930968761444092
Iter #757760:  Learning rate = 0.003757:   Batch Loss = 1.193168, Accuracy
= 0.779296875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2815539836883545, Accu
racy = 0.7308294177055359
Iter #761856:  Learning rate = 0.003757:   Batch Loss = 1.159685, Accuracy
= 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2689461708068848, Accu
racy = 0.7330899238586426
Iter #765952:  Learning rate = 0.003757:   Batch Loss = 1.178314, Accuracy
= 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2470251321792603, Accu
racy = 0.7356981635093689
Iter #770048:  Learning rate = 0.003757:   Batch Loss = 1.221081, Accuracy
= 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.294454574584961, Accur
acy = 0.7080507874488831
Iter #774144:  Learning rate = 0.003757:   Batch Loss = 1.257118, Accuracy
= 0.73046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2873599529266357, Accu
racy = 0.688401997089386
Iter #778240:  Learning rate = 0.003757:   Batch Loss = 1.226128, Accuracy
= 0.74609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2573223114013672, Accu
racy = 0.7303077578544617
Iter #782336:  Learning rate = 0.003757:   Batch Loss = 1.239028, Accuracy
= 0.73046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2715098857879639, Accu
racy = 0.6997044086456299
Iter #786432:  Learning rate = 0.003757:   Batch Loss = 1.204247, Accuracy
= 0.751953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2909232378005981, Accu
racy = 0.7280473113059998
Iter #790528:  Learning rate = 0.003757:   Batch Loss = 1.178913, Accuracy
= 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2500776052474976, Accu
racy = 0.7355242371559143
Iter #794624:  Learning rate = 0.003757:   Batch Loss = 1.234663, Accuracy
= 0.744140625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2899547815322876, Accu
racy = 0.6863154172897339
Iter #798720:  Learning rate = 0.003757:   Batch Loss = 1.186783, Accuracy
= 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.249901533126831, Accur
acy = 0.7323943376541138
Iter #802816:  Learning rate = 0.003607:   Batch Loss = 1.245442, Accuracy
= 0.724609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2630853652954102, Accu
racy = 0.7287428379058838
Iter #806912:  Learning rate = 0.003607:   Batch Loss = 1.240891, Accuracy
= 0.7265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2730265855789185, Accu
racy = 0.6924013495445251
Iter #811008:  Learning rate = 0.003607:   Batch Loss = 1.246924, Accuracy
```

```
= 0.728515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2656391859054565, Accu
racy = 0.7407407164573669
Iter #815104:  Learning rate = 0.003607:   Batch Loss = 1.155763, Accuracy
= 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2334394454956055, Accu
racy = 0.7416101694107056
Iter #819200:  Learning rate = 0.003607:   Batch Loss = 1.270316, Accuracy
= 0.70703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2428205013275146, Accu
racy = 0.7358720302581787
Iter #823296:  Learning rate = 0.003607:   Batch Loss = 1.194757, Accuracy
= 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 1.23284912109375, Accura
cy = 0.7388280034065247
Iter #827392:  Learning rate = 0.003607:   Batch Loss = 1.212878, Accuracy
= 0.73828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.3371522426605225, Accu
racy = 0.6697965860366821
Iter #831488:  Learning rate = 0.003607:   Batch Loss = 1.219350, Accuracy
= 0.744140625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2593777179718018, Accu
racy = 0.7308294177055359
Iter #835584:  Learning rate = 0.003607:   Batch Loss = 1.214573, Accuracy
= 0.728515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.31206214427948, Accura
cy = 0.7050947546958923
Iter #839680:  Learning rate = 0.003607:   Batch Loss = 1.233889, Accuracy
= 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.263655424118042, Accur
acy = 0.724395751953125
Iter #843776:  Learning rate = 0.003607:   Batch Loss = 1.263289, Accuracy
= 0.712890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2768659591674805, Accu
racy = 0.7316988110542297
Iter #847872:  Learning rate = 0.003607:   Batch Loss = 1.207096, Accuracy
= 0.73828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2700845003128052, Accu
racy = 0.7264823317527771
Iter #851968:  Learning rate = 0.003607:   Batch Loss = 1.202019, Accuracy
= 0.73828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.255967617034912, Accur
acy = 0.7374369502067566
Iter #856064:  Learning rate = 0.003607:   Batch Loss = 1.223450, Accuracy
= 0.71875
PERFORMANCE ON TEST SET:             Batch Loss = 1.269404411315918, Accur
acy = 0.6924013495445251
Iter #860160:  Learning rate = 0.003607:   Batch Loss = 1.172800, Accuracy
= 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2376048564910889, Accu
racy = 0.7431750893592834
Iter #864256:  Learning rate = 0.003607:   Batch Loss = 1.122565, Accuracy
= 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2062935829162598, Accu
racy = 0.7509998083114624
Iter #868352:  Learning rate = 0.003607:   Batch Loss = 1.277532, Accuracy
= 0.7265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2447361946105957, Accu
racy = 0.7332637906074524
Iter #872448:  Learning rate = 0.003607:   Batch Loss = 1.190871, Accuracy
= 0.74609375
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 1.2701406478881836, Accu
racy = 0.6804034113883972
Iter #876544:  Learning rate = 0.003607:   Batch Loss = 1.221617, Accuracy
= 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2214765548706055, Accu
racy = 0.7405668497085571
Iter #880640:  Learning rate = 0.003607:   Batch Loss = 1.181198, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2424073219299316, Accu
racy = 0.7356981635093689
Iter #884736:  Learning rate = 0.003607:   Batch Loss = 1.266841, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.3276904821395874, Accu
racy = 0.7089201807975769
Iter #888832:  Learning rate = 0.003607:   Batch Loss = 1.315230, Accuracy
= 0.716796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.4555960893630981, Accu
racy = 0.665797233581543
Iter #892928:  Learning rate = 0.003607:   Batch Loss = 1.260028, Accuracy
= 0.720703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2883548736572266, Accu
racy = 0.7101373672485352
Iter #897024:  Learning rate = 0.003607:   Batch Loss = 1.191475, Accuracy
= 0.724609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.209590196609497, Accur
acy = 0.7235263586044312
Iter #901120:  Learning rate = 0.003463:   Batch Loss = 1.223314, Accuracy
= 0.75
PERFORMANCE ON TEST SET:             Batch Loss = 1.269482135772705, Accur
acy = 0.7334376573562622
Iter #905216:  Learning rate = 0.003463:   Batch Loss = 1.359107, Accuracy
= 0.693359375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2801330089569092, Accu
racy = 0.7271778583526611
Iter #909312:  Learning rate = 0.003463:   Batch Loss = 1.135830, Accuracy
= 0.75390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2292490005493164, Accu
racy = 0.7398713231086731
Iter #913408:  Learning rate = 0.003463:   Batch Loss = 1.206129, Accuracy
= 0.751953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2572212219238281, Accu
racy = 0.7290905714035034
Iter #917504:  Learning rate = 0.003463:   Batch Loss = 1.106161, Accuracy
= 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.20521879196167, Accura
cy = 0.7457833290100098
Iter #921600:  Learning rate = 0.003463:   Batch Loss = 1.075208, Accuracy
= 0.80078125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2007334232330322, Accu
racy = 0.7143105268478394
Iter #925696:  Learning rate = 0.003463:   Batch Loss = 1.165758, Accuracy
= 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2181956768035889, Accu
racy = 0.7400451898574829
Iter #929792:  Learning rate = 0.003463:   Batch Loss = 1.123657, Accuracy
= 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.226883888244629, Accur
acy = 0.7271778583526611
Iter #933888:  Learning rate = 0.003463:   Batch Loss = 1.153775, Accuracy
= 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2505993843078613, Accu
```

```
racy = 0.693270742893219
Iter #937984:  Learning rate = 0.003463:   Batch Loss = 1.099388, Accuracy
= 0.78125
PERFORMANCE ON TEST SET:             Batch Loss = 1.205772876739502, Accur
acy = 0.7099635004997253
Iter #942080:  Learning rate = 0.003463:   Batch Loss = 1.213825, Accuracy
= 0.7109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1982582807540894, Accu
racy = 0.7475221753120422
Iter #946176:  Learning rate = 0.003463:   Batch Loss = 1.238396, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1784837245941162, Accu
racy = 0.7501304149627686
Iter #950272:  Learning rate = 0.003463:   Batch Loss = 1.205057, Accuracy
= 0.740234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2215361595153809, Accu
racy = 0.7421318292617798
Iter #954368:  Learning rate = 0.003463:   Batch Loss = 1.204586, Accuracy
= 0.728515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2257506847381592, Accu
racy = 0.7157016396522522
Iter #958464:  Learning rate = 0.003463:   Batch Loss = 1.162105, Accuracy
= 0.748046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2051020860671997, Accu
racy = 0.7435228824615479
Iter #962560:  Learning rate = 0.003463:   Batch Loss = 1.157220, Accuracy
= 0.75
PERFORMANCE ON TEST SET:             Batch Loss = 1.2265043258666992, Accu
racy = 0.7350026369094849
Iter #966656:  Learning rate = 0.003463:   Batch Loss = 1.113189, Accuracy
= 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2094910144805908, Accu
racy = 0.737784743309021
Iter #970752:  Learning rate = 0.003463:   Batch Loss = 1.109542, Accuracy
= 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.205865502357483, Accur
acy = 0.7104851603507996
Iter #974848:  Learning rate = 0.003463:   Batch Loss = 1.151349, Accuracy
= 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1983461380004883, Accu
racy = 0.7403929829597473
Iter #978944:  Learning rate = 0.003463:   Batch Loss = 1.237178, Accuracy
= 0.7265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1944518089294434, Accu
racy = 0.7423056960105896
Iter #983040:  Learning rate = 0.003463:   Batch Loss = 1.215197, Accuracy
= 0.732421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2176823616027832, Accu
racy = 0.7155277132987976
Iter #987136:  Learning rate = 0.003463:   Batch Loss = 1.138189, Accuracy
= 0.75390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2076005935668945, Accu
racy = 0.7449139356613159
Iter #991232:  Learning rate = 0.003463:   Batch Loss = 1.242819, Accuracy
= 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.22664213180542, Accura
cy = 0.737089216709137
Iter #995328:  Learning rate = 0.003463:   Batch Loss = 1.152296, Accuracy
= 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2383599281311035, Accu
racy = 0.7339593172073364
```

```
Iter #999424:  Learning rate = 0.003463:   Batch Loss = 1.173647, Accuracy
= 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2274386882781982, Accu
racy = 0.7320466041564941
Iter #1003520:  Learning rate = 0.003324:   Batch Loss = 1.148685, Accurac
y = 0.75
PERFORMANCE ON TEST SET:             Batch Loss = 1.2099651098251343, Accu
racy = 0.6988349556922913
Iter #1007616:  Learning rate = 0.003324:   Batch Loss = 1.110898, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2090102434158325, Accu
racy = 0.7372630834579468
Iter #1011712:  Learning rate = 0.003324:   Batch Loss = 1.146289, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2187511920928955, Accu
racy = 0.6927490830421448
Iter #1015808:  Learning rate = 0.003324:   Batch Loss = 1.139839, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2025260925292969, Accu
racy = 0.7442184090614319
Iter #1019904:  Learning rate = 0.003324:   Batch Loss = 1.174099, Accurac
y = 0.740234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.178891897201538, Accur
acy = 0.7489132285118103
Iter #1024000:  Learning rate = 0.003324:   Batch Loss = 1.126517, Accurac
y = 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1835594177246094, Accu
racy = 0.7457833290100098
Iter #1028096:  Learning rate = 0.003324:   Batch Loss = 1.133618, Accurac
y = 0.73828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1934621334075928, Accu
racy = 0.7172665596008301
Iter #1032192:  Learning rate = 0.003324:   Batch Loss = 1.195170, Accurac
y = 0.724609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.176482915878296, Accur
acy = 0.7235263586044312
Iter #1036288:  Learning rate = 0.003324:   Batch Loss = 1.138812, Accurac
y = 0.734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2366979122161865, Accu
racy = 0.7313510775566101
Iter #1040384:  Learning rate = 0.003324:   Batch Loss = 1.100087, Accurac
y = 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2413005828857422, Accu
racy = 0.7296122312545776
Iter #1044480:  Learning rate = 0.003324:   Batch Loss = 1.163728, Accurac
y = 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2169625759124756, Accu
racy = 0.7325682640075684
Iter #1048576:  Learning rate = 0.003324:   Batch Loss = 1.239370, Accurac
y = 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2068835496902466, Accu
racy = 0.7332637906074524
Iter #1052672:  Learning rate = 0.003324:   Batch Loss = 1.147866, Accurac
y = 0.75390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2324965000152588, Accu
racy = 0.7313510775566101
Iter #1056768:  Learning rate = 0.003324:   Batch Loss = 1.110332, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1922078132629395, Accu
racy = 0.7473483085632324
Iter #1060864:  Learning rate = 0.003324:   Batch Loss = 1.219769, Accurac
```

```
y = 0.6875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1736798286437988, Accu
racy = 0.7490870952606201
Iter #1064960:  Learning rate = 0.003324:   Batch Loss = 1.068718, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.182934045791626, Accur
acy = 0.7407407164573669
Iter #1069056:  Learning rate = 0.003324:   Batch Loss = 1.122719, Accurac
y = 0.75390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1685690879821777, Accu
racy = 0.747696042060852
Iter #1073152:  Learning rate = 0.003324:   Batch Loss = 1.141710, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2125314474105835, Accu
racy = 0.6910102367401123
Iter #1077248:  Learning rate = 0.003324:   Batch Loss = 1.191853, Accurac
y = 0.732421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2111785411834717, Accu
racy = 0.7395235896110535
Iter #1081344:  Learning rate = 0.003324:   Batch Loss = 1.144769, Accurac
y = 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2182832956314087, Accu
racy = 0.7155277132987976
Iter #1085440:  Learning rate = 0.003324:   Batch Loss = 1.178873, Accurac
y = 0.74609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2005038261413574, Accu
racy = 0.7438706159591675
Iter #1089536:  Learning rate = 0.003324:   Batch Loss = 1.180037, Accurac
y = 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.207581877708435, Accur
acy = 0.7363936901092529
Iter #1093632:  Learning rate = 0.003324:   Batch Loss = 1.131612, Accurac
y = 0.751953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1856026649475098, Accu
racy = 0.7442184090614319
Iter #1097728:  Learning rate = 0.003324:   Batch Loss = 1.162144, Accurac
y = 0.728515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1758962869644165, Accu
racy = 0.7468266487121582
Iter #1101824:  Learning rate = 0.003191:   Batch Loss = 1.100184, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.161110281944275, Accur
acy = 0.7577812671661377
Iter #1105920:  Learning rate = 0.003191:   Batch Loss = 1.072707, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1880135536193848, Accu
racy = 0.7146583199501038
Iter #1110016:  Learning rate = 0.003191:   Batch Loss = 1.123458, Accurac
y = 0.748046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1687402725219727, Accu
racy = 0.7497826218605042
Iter #1114112:  Learning rate = 0.003191:   Batch Loss = 1.104726, Accurac
y = 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1677427291870117, Accu
racy = 0.7468266487121582
Iter #1118208:  Learning rate = 0.003191:   Batch Loss = 1.117300, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.240354061126709, Accur
acy = 0.7273517847061157
Iter #1122304:  Learning rate = 0.003191:   Batch Loss = 1.162065, Accurac
y = 0.740234375
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 1.1769474744796753, Accu
racy = 0.7186576128005981
Iter #1126400:  Learning rate = 0.003191:   Batch Loss = 1.103808, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1816388368606567, Accu
racy = 0.7360458970069885
Iter #1130496:  Learning rate = 0.003191:   Batch Loss = 1.113840, Accurac
y = 0.740234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2219798564910889, Accu
racy = 0.6974439024925232
Iter #1134592:  Learning rate = 0.003191:   Batch Loss = 1.177770, Accurac
y = 0.712890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1869109869003296, Accu
racy = 0.7383064031600952
Iter #1138688:  Learning rate = 0.003191:   Batch Loss = 1.094514, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1651204824447632, Accu
racy = 0.7527386546134949
Iter #1142784:  Learning rate = 0.003191:   Batch Loss = 1.080745, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1697293519973755, Accu
racy = 0.7452616691589355
Iter #1146880:  Learning rate = 0.003191:   Batch Loss = 1.099609, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1377360820770264, Accu
racy = 0.758998453617096
Iter #1150976:  Learning rate = 0.003191:   Batch Loss = 1.093319, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1396219730377197, Accu
racy = 0.754825234413147
Iter #1155072:  Learning rate = 0.003191:   Batch Loss = 1.036572, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1371163129806519, Accu
racy = 0.7355242371559143
Iter #1159168:  Learning rate = 0.003191:   Batch Loss = 1.063365, Accurac
y = 0.78125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1456021070480347, Accu
racy = 0.7520431280136108
Iter #1163264:  Learning rate = 0.003191:   Batch Loss = 1.099315, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.150299310684204, Accur
acy = 0.7509998083114624
Iter #1167360:  Learning rate = 0.003191:   Batch Loss = 1.044715, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.124154806137085, Accur
acy = 0.7600417137145996
Iter #1171456:  Learning rate = 0.003191:   Batch Loss = 1.079661, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1579296588897705, Accu
racy = 0.7530864477157593
Iter #1175552:  Learning rate = 0.003191:   Batch Loss = 1.151993, Accurac
y = 0.740234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2518888711929321, Accu
racy = 0.7104851603507996
Iter #1179648:  Learning rate = 0.003191:   Batch Loss = 1.137970, Accurac
y = 0.75390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2066452503204346, Accu
racy = 0.7337854504585266
Iter #1183744:  Learning rate = 0.003191:   Batch Loss = 1.256392, Accurac
y = 0.685546875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1551403999328613, Accu
```

```
racy = 0.7520431280136108
Iter #1187840:  Learning rate = 0.003191:   Batch Loss = 1.118422, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1554440259933472, Accu
racy = 0.7475221753120422
Iter #1191936:  Learning rate = 0.003191:   Batch Loss = 1.035783, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1319046020507812, Accu
racy = 0.7558685541152954
Iter #1196032:  Learning rate = 0.003191:   Batch Loss = 1.110743, Accurac
y = 0.748046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1277339458465576, Accu
racy = 0.7595200538635254
Iter #1200128:  Learning rate = 0.003064:   Batch Loss = 1.190066, Accurac
y = 0.740234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.2948157787322998, Accu
racy = 0.6990088820457458
Iter #1204224:  Learning rate = 0.003064:   Batch Loss = 1.020924, Accurac
y = 0.80859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1926498413085938, Accu
racy = 0.7310032844543457
Iter #1208320:  Learning rate = 0.003064:   Batch Loss = 1.195525, Accurac
y = 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1297305822372437, Accu
racy = 0.7605633735656738
Iter #1212416:  Learning rate = 0.003064:   Batch Loss = 1.094507, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1512718200683594, Accu
racy = 0.7541297078132629
Iter #1216512:  Learning rate = 0.003064:   Batch Loss = 1.137871, Accurac
y = 0.740234375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1082228422164917, Accu
racy = 0.7680403590202332
Iter #1220608:  Learning rate = 0.003064:   Batch Loss = 1.045623, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1285738945007324, Accu
racy = 0.7544775009155273
Iter #1224704:  Learning rate = 0.003064:   Batch Loss = 1.058982, Accurac
y = 0.796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1034573316574097, Accu
racy = 0.7591723203659058
Iter #1228800:  Learning rate = 0.003064:   Batch Loss = 1.129120, Accurac
y = 0.75390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1381762027740479, Accu
racy = 0.7563902139663696
Iter #1232896:  Learning rate = 0.003064:   Batch Loss = 1.048266, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1590545177459717, Accu
racy = 0.7226569056510925
Iter #1236992:  Learning rate = 0.003064:   Batch Loss = 1.081789, Accurac
y = 0.748046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1505863666534424, Accu
racy = 0.7202225923538208
Iter #1241088:  Learning rate = 0.003064:   Batch Loss = 1.071091, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1201765537261963, Accu
racy = 0.7642149329185486
Iter #1245184:  Learning rate = 0.003064:   Batch Loss = 0.987554, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.089844822883606, Accur
acy = 0.7381324768066406
```

```
Iter #1249280:  Learning rate = 0.003064:   Batch Loss = 1.088638, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1103335618972778, Accu
racy = 0.7572596073150635
Iter #1253376:  Learning rate = 0.003064:   Batch Loss = 1.086179, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1209150552749634, Accu
racy = 0.7650843262672424
Iter #1257472:  Learning rate = 0.003064:   Batch Loss = 1.086447, Accurac
y = 0.76171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1016472578048706, Accu
racy = 0.7638671398162842
Iter #1261568:  Learning rate = 0.003064:   Batch Loss = 1.031532, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0916762351989746, Accu
racy = 0.7697791457176208
Iter #1265664:  Learning rate = 0.003064:   Batch Loss = 1.085721, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1414904594421387, Accu
racy = 0.7485654950141907
Iter #1269760:  Learning rate = 0.003064:   Batch Loss = 1.082261, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.140390157699585, Accur
acy = 0.7259607315063477
Iter #1273856:  Learning rate = 0.003064:   Batch Loss = 1.007086, Accurac
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.117671251296997, Accur
acy = 0.76334547996521
Iter #1277952:  Learning rate = 0.003064:   Batch Loss = 0.992834, Accurac
y = 0.810546875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0951639413833618, Accu
racy = 0.7697791457176208
Iter #1282048:  Learning rate = 0.003064:   Batch Loss = 1.059710, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1145819425582886, Accu
racy = 0.7600417137145996
Iter #1286144:  Learning rate = 0.003064:   Batch Loss = 1.050372, Accurac
y = 0.80078125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1324446201324463, Accu
racy = 0.755520761013031
Iter #1290240:  Learning rate = 0.003064:   Batch Loss = 1.057630, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1175663471221924, Accu
racy = 0.75969398021698
Iter #1294336:  Learning rate = 0.003064:   Batch Loss = 1.184515, Accurac
y = 0.716796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1471482515335083, Accu
racy = 0.7522169947624207
Iter #1298432:  Learning rate = 0.003064:   Batch Loss = 1.031288, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0969882011413574, Accu
racy = 0.7405668497085571
Iter #1302528:  Learning rate = 0.002941:   Batch Loss = 1.019033, Accurac
y = 0.78125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0974323749542236, Accu
racy = 0.7638671398162842
Iter #1306624:  Learning rate = 0.002941:   Batch Loss = 1.135507, Accurac
y = 0.748046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1332731246948242, Accu
racy = 0.7276995182037354
Iter #1310720:  Learning rate = 0.002941:   Batch Loss = 1.025180, Accurac
```

```
y = 0.78125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1182600259780884, Accu
racy = 0.7690836191177368
Iter #1314816:  Learning rate = 0.002941:   Batch Loss = 1.136992, Accurac
y = 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1834638118743896, Accu
racy = 0.7350026369094849
Iter #1318912:  Learning rate = 0.002941:   Batch Loss = 1.045726, Accurac
y = 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 1.143284559249878, Accur
acy = 0.7195270657539368
Iter #1323008:  Learning rate = 0.002941:   Batch Loss = 1.015148, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0645220279693604, Accu
racy = 0.7753434181213379
Iter #1327104:  Learning rate = 0.002941:   Batch Loss = 1.069362, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0790313482284546, Accu
racy = 0.7749956250190735
Iter #1331200:  Learning rate = 0.002941:   Batch Loss = 1.077284, Accurac
y = 0.75390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0978786945343018, Accu
racy = 0.7344809770584106
Iter #1335296:  Learning rate = 0.002941:   Batch Loss = 1.104882, Accurac
y = 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.2542698383331299, Accu
racy = 0.6948356628417969
Iter #1339392:  Learning rate = 0.002941:   Batch Loss = 1.033936, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1160354614257812, Accu
racy = 0.7556946873664856
Iter #1343488:  Learning rate = 0.002941:   Batch Loss = 1.088041, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.122750163078308, Accur
acy = 0.7586506605148315
Iter #1347584:  Learning rate = 0.002941:   Batch Loss = 1.052256, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1215603351593018, Accu
racy = 0.7308294177055359
Iter #1351680:  Learning rate = 0.002941:   Batch Loss = 1.065774, Accurac
y = 0.76171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0660760402679443, Accu
racy = 0.7763867378234863
Iter #1355776:  Learning rate = 0.002941:   Batch Loss = 1.001461, Accurac
y = 0.796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0763602256774902, Accu
racy = 0.7770822644233704
Iter #1359872:  Learning rate = 0.002941:   Batch Loss = 0.980563, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.069732666015625, Accur
acy = 0.7494348883628845
Iter #1363968:  Learning rate = 0.002941:   Batch Loss = 1.004383, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0604956150054932, Accu
racy = 0.755520761013031
Iter #1368064:  Learning rate = 0.002941:   Batch Loss = 1.037492, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0756494998931885, Accu
racy = 0.7386541366577148
Iter #1372160:  Learning rate = 0.002941:   Batch Loss = 0.986197, Accurac
y = 0.80078125
```

```
PERFORMANCE ON TEST SET:            Batch Loss = 1.0785447359085083, Accu
racy = 0.7642149329185486
Iter #1376256:  Learning rate = 0.002941:   Batch Loss = 0.999449, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:            Batch Loss = 1.0671353340148926, Accu
racy = 0.7715179920196533
Iter #1380352:  Learning rate = 0.002941:   Batch Loss = 1.094628, Accurac
y = 0.75390625
PERFORMANCE ON TEST SET:            Batch Loss = 1.129581093788147, Accur
acy = 0.7198747992515564
Iter #1384448:  Learning rate = 0.002941:   Batch Loss = 1.018920, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:            Batch Loss = 1.0830916166305542, Accu
racy = 0.769605278968811
Iter #1388544:  Learning rate = 0.002941:   Batch Loss = 0.982512, Accurac
y = 0.814453125
PERFORMANCE ON TEST SET:            Batch Loss = 1.0939812660217285, Accu
racy = 0.7450878024101257
Iter #1392640:  Learning rate = 0.002941:   Batch Loss = 1.086622, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:            Batch Loss = 1.0775742530822754, Accu
racy = 0.7706485986709595
Iter #1396736:  Learning rate = 0.002941:   Batch Loss = 1.011496, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:            Batch Loss = 1.0529260635375977, Accu
racy = 0.7795165777206421
Iter #1400832:  Learning rate = 0.002823:   Batch Loss = 1.081436, Accurac
y = 0.75390625
PERFORMANCE ON TEST SET:            Batch Loss = 1.0997284650802612, Accu
racy = 0.7323943376541138
Iter #1404928:  Learning rate = 0.002823:   Batch Loss = 1.030347, Accurac
y = 0.78125
PERFORMANCE ON TEST SET:            Batch Loss = 1.1064238548278809, Accu
racy = 0.741436243057251
Iter #1409024:  Learning rate = 0.002823:   Batch Loss = 1.029403, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:            Batch Loss = 1.0846680402755737, Accu
racy = 0.7443922758102417
Iter #1413120:  Learning rate = 0.002823:   Batch Loss = 1.060438, Accurac
y = 0.76171875
PERFORMANCE ON TEST SET:            Batch Loss = 1.2092645168304443, Accu
racy = 0.7163971662521362
Iter #1417216:  Learning rate = 0.002823:   Batch Loss = 1.101009, Accurac
y = 0.75
PERFORMANCE ON TEST SET:            Batch Loss = 1.132475733757019, Accur
acy = 0.7410885095596313
Iter #1421312:  Learning rate = 0.002823:   Batch Loss = 1.062394, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:            Batch Loss = 1.052087426185608, Accur
acy = 0.7553468942642212
Iter #1425408:  Learning rate = 0.002823:   Batch Loss = 1.042004, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:            Batch Loss = 1.2360265254974365, Accu
racy = 0.6892714500427246
Iter #1429504:  Learning rate = 0.002823:   Batch Loss = 1.010729, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:            Batch Loss = 1.0981879234313965, Accu
racy = 0.7628238797187805
Iter #1433600:  Learning rate = 0.002823:   Batch Loss = 0.975663, Accurac
y = 0.80078125
PERFORMANCE ON TEST SET:            Batch Loss = 1.1192209720611572, Accu
```

```
racy = 0.7593461871147156
Iter #1437696:  Learning rate = 0.002823:   Batch Loss = 1.117724, Accurac
y = 0.755859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0803489685058594, Accu
racy = 0.7659537196159363
Iter #1441792:  Learning rate = 0.002823:   Batch Loss = 0.988126, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1528129577636719, Accu
racy = 0.7351765036582947
Iter #1445888:  Learning rate = 0.002823:   Batch Loss = 1.018625, Accurac
y = 0.78515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0774723291397095, Accu
racy = 0.7725613117218018
Iter #1449984:  Learning rate = 0.002823:   Batch Loss = 1.023714, Accurac
y = 0.796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.3513617515563965, Accu
racy = 0.6817944645881653
Iter #1454080:  Learning rate = 0.002823:   Batch Loss = 1.002318, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0869040489196777, Accu
racy = 0.7642149329185486
Iter #1458176:  Learning rate = 0.002823:   Batch Loss = 1.077110, Accurac
y = 0.734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.092512607574463, Accur
acy = 0.7395235896110535
Iter #1462272:  Learning rate = 0.002823:   Batch Loss = 0.989524, Accurac
y = 0.810546875
PERFORMANCE ON TEST SET:             Batch Loss = 1.058807373046875, Accur
acy = 0.7739523649215698
Iter #1466368:  Learning rate = 0.002823:   Batch Loss = 1.035892, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0597398281097412, Accu
racy = 0.7716918587684631
Iter #1470464:  Learning rate = 0.002823:   Batch Loss = 0.999967, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.047798752784729, Accur
acy = 0.7784733176231384
Iter #1474560:  Learning rate = 0.002823:   Batch Loss = 0.988230, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0826596021652222, Accu
racy = 0.7393496632575989
Iter #1478656:  Learning rate = 0.002823:   Batch Loss = 0.968140, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0683870315551758, Accu
racy = 0.7426534295082092
Iter #1482752:  Learning rate = 0.002823:   Batch Loss = 1.021802, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.06016206741333, Accura
cy = 0.7704746723175049
Iter #1486848:  Learning rate = 0.002823:   Batch Loss = 1.002370, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0800378322601318, Accu
racy = 0.747696042060852
Iter #1490944:  Learning rate = 0.002823:   Batch Loss = 0.975460, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0677701234817505, Accu
racy = 0.7543035745620728
Iter #1495040:  Learning rate = 0.002823:   Batch Loss = 0.989045, Accurac
y = 0.806640625
PERFORMANCE ON TEST SET:             Batch Loss = 1.032030463218689, Accur
acy = 0.778125524520874
```

```
Iter #1499136:  Learning rate = 0.002823:   Batch Loss = 0.920823, Accurac
y = 0.833984375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0400888919830322, Accu
racy = 0.7782994508743286
Iter #1503232:  Learning rate = 0.002710:   Batch Loss = 1.084355, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0827560424804688, Accu
racy = 0.7525647878646851
Iter #1507328:  Learning rate = 0.002710:   Batch Loss = 1.004571, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0604846477508545, Accu
racy = 0.7657798528671265
Iter #1511424:  Learning rate = 0.002710:   Batch Loss = 0.961999, Accurac
y = 0.80859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0682778358459473, Accu
racy = 0.7504781484603882
Iter #1515520:  Learning rate = 0.002710:   Batch Loss = 0.963803, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0499320030212402, Accu
racy = 0.7537819743156433
Iter #1519616:  Learning rate = 0.002710:   Batch Loss = 1.036237, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0532357692718506, Accu
racy = 0.7504781484603882
Iter #1523712:  Learning rate = 0.002710:   Batch Loss = 1.083717, Accurac
y = 0.748046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0663869380950928, Accu
racy = 0.7428273558616638
Iter #1527808:  Learning rate = 0.002710:   Batch Loss = 1.023219, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0791935920715332, Accu
racy = 0.7711702585220337
Iter #1531904:  Learning rate = 0.002710:   Batch Loss = 0.954356, Accurac
y = 0.806640625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0383235216140747, Accu
racy = 0.7626499533653259
Iter #1536000:  Learning rate = 0.002710:   Batch Loss = 0.963677, Accurac
y = 0.79296875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0435419082641602, Accu
racy = 0.7569118142127991
Iter #1540096:  Learning rate = 0.002710:   Batch Loss = 1.000395, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0436553955078125, Accu
racy = 0.7541297078132629
Iter #1544192:  Learning rate = 0.002710:   Batch Loss = 1.022665, Accurac
y = 0.779296875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0347216129302979, Accu
racy = 0.7494348883628845
Iter #1548288:  Learning rate = 0.002710:   Batch Loss = 1.028026, Accurac
y = 0.78515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0825071334838867, Accu
racy = 0.764041006565094
Iter #1552384:  Learning rate = 0.002710:   Batch Loss = 0.992770, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.053978443145752, Accur
acy = 0.7499565482139587
Iter #1556480:  Learning rate = 0.002710:   Batch Loss = 0.950835, Accurac
y = 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0554957389831543, Accu
racy = 0.7449139356613159
Iter #1560576:  Learning rate = 0.002710:   Batch Loss = 1.016983, Accurac
```

```
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0589797496795654, Accu
racy = 0.761085033416748
Iter #1564672:  Learning rate = 0.002710:   Batch Loss = 1.006269, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0475605726242065, Accu
racy = 0.7814292907714844
Iter #1568768:  Learning rate = 0.002710:   Batch Loss = 1.010420, Accurac
y = 0.78125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0468571186065674, Accu
racy = 0.773082971572876
Iter #1572864:  Learning rate = 0.002710:   Batch Loss = 0.980588, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.097536563873291, Accur
acy = 0.7654321193695068
Iter #1576960:  Learning rate = 0.002710:   Batch Loss = 0.981525, Accurac
y = 0.8203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0341451168060303, Accu
racy = 0.77377849817276
Iter #1581056:  Learning rate = 0.002710:   Batch Loss = 1.016225, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0494534969329834, Accu
racy = 0.774474024772644
Iter #1585152:  Learning rate = 0.002710:   Batch Loss = 1.031667, Accurac
y = 0.76171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0799517631530762, Accu
racy = 0.7445661425590515
Iter #1589248:  Learning rate = 0.002710:   Batch Loss = 0.965658, Accurac
y = 0.78515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0403631925582886, Accu
racy = 0.7791688442230225
Iter #1593344:  Learning rate = 0.002710:   Batch Loss = 0.943323, Accurac
y = 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 1.064902663230896, Accur
acy = 0.7480438351631165
Iter #1597440:  Learning rate = 0.002710:   Batch Loss = 1.009954, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0514155626296997, Accu
racy = 0.7543035745620728
Iter #1601536:  Learning rate = 0.002602:   Batch Loss = 0.947811, Accurac
y = 0.806640625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0739537477493286, Accu
racy = 0.7569118142127991
Iter #1605632:  Learning rate = 0.002602:   Batch Loss = 1.041207, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1098374128341675, Accu
racy = 0.7221353054046631
Iter #1609728:  Learning rate = 0.002602:   Batch Loss = 1.055181, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0811060667037964, Accu
racy = 0.7636932730674744
Iter #1613824:  Learning rate = 0.002602:   Batch Loss = 0.984855, Accurac
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0700128078460693, Accu
racy = 0.7579551339149475
Iter #1617920:  Learning rate = 0.002602:   Batch Loss = 1.017498, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.066851019859314, Accur
acy = 0.7614327669143677
Iter #1622016:  Learning rate = 0.002602:   Batch Loss = 0.976790, Accurac
y = 0.78515625
```

```
PERFORMANCE ON TEST SET:            Batch Loss = 1.0506854057312012, Accu
racy = 0.7709963321685791
Iter #1626112:  Learning rate = 0.002602:   Batch Loss = 1.047369, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:            Batch Loss = 1.047089695930481, Accur
acy = 0.7708224654197693
Iter #1630208:  Learning rate = 0.002602:   Batch Loss = 1.013215, Accurac
y = 0.78515625
PERFORMANCE ON TEST SET:            Batch Loss = 1.09821617603302, Accura
cy = 0.7433489561080933
Iter #1634304:  Learning rate = 0.002602:   Batch Loss = 0.913662, Accurac
y = 0.826171875
PERFORMANCE ON TEST SET:            Batch Loss = 1.0516241788864136, Accu
racy = 0.7758650779724121
Iter #1638400:  Learning rate = 0.002602:   Batch Loss = 1.098814, Accurac
y = 0.744140625
PERFORMANCE ON TEST SET:            Batch Loss = 1.1114096641540527, Accu
racy = 0.7191792726516724
Iter #1642496:  Learning rate = 0.002602:   Batch Loss = 1.005033, Accurac
y = 0.78125
PERFORMANCE ON TEST SET:            Batch Loss = 1.065004587173462, Accur
acy = 0.7703008055686951
Iter #1646592:  Learning rate = 0.002602:   Batch Loss = 0.975548, Accurac
y = 0.796875
PERFORMANCE ON TEST SET:            Batch Loss = 1.0457868576049805, Accu
racy = 0.7471743822097778
Iter #1650688:  Learning rate = 0.002602:   Batch Loss = 1.021883, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:            Batch Loss = 1.0710010528564453, Accu
racy = 0.7668231725692749
Iter #1654784:  Learning rate = 0.002602:   Batch Loss = 1.012290, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:            Batch Loss = 1.0645990371704102, Accu
racy = 0.7358720302581787
Iter #1658880:  Learning rate = 0.002602:   Batch Loss = 0.964328, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:            Batch Loss = 1.0475435256958008, Accu
racy = 0.7629977464675903
Iter #1662976:  Learning rate = 0.002602:   Batch Loss = 1.023030, Accurac
y = 0.796875
PERFORMANCE ON TEST SET:            Batch Loss = 1.0411654710769653, Accu
racy = 0.7489132285118103
Iter #1667072:  Learning rate = 0.002602:   Batch Loss = 1.015789, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:            Batch Loss = 1.052968978881836, Accur
acy = 0.7706485986709595
Iter #1671168:  Learning rate = 0.002602:   Batch Loss = 0.992228, Accurac
y = 0.79296875
PERFORMANCE ON TEST SET:            Batch Loss = 1.0453262329101562, Accu
racy = 0.7675186991691589
Iter #1675264:  Learning rate = 0.002602:   Batch Loss = 1.006626, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:            Batch Loss = 1.0343981981277466, Accu
racy = 0.7736045718193054
Iter #1679360:  Learning rate = 0.002602:   Batch Loss = 1.134959, Accurac
y = 0.712890625
PERFORMANCE ON TEST SET:            Batch Loss = 1.0683435201644897, Accu
racy = 0.734307050704956
Iter #1683456:  Learning rate = 0.002602:   Batch Loss = 1.029807, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:            Batch Loss = 1.108661413192749, Accur
```

```
acy = 0.7501304149627686
Iter #1687552:  Learning rate = 0.002602:   Batch Loss = 1.002633, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0694506168365479, Accu
racy = 0.7680403590202332
Iter #1691648:  Learning rate = 0.002602:   Batch Loss = 0.931768, Accurac
y = 0.8046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.074845790863037, Accur
acy = 0.7593461871147156
Iter #1695744:  Learning rate = 0.002602:   Batch Loss = 0.987920, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0363272428512573, Accu
racy = 0.7729090452194214
Iter #1699840:  Learning rate = 0.002602:   Batch Loss = 0.996746, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0523463487625122, Accu
racy = 0.7664753794670105
Iter #1703936:  Learning rate = 0.002498:   Batch Loss = 0.943786, Accurac
y = 0.80859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0273220539093018, Accu
racy = 0.7546513676643372
Iter #1708032:  Learning rate = 0.002498:   Batch Loss = 1.014014, Accurac
y = 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0647377967834473, Accu
racy = 0.7296122312545776
Iter #1712128:  Learning rate = 0.002498:   Batch Loss = 0.925235, Accurac
y = 0.814453125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0534108877182007, Accu
racy = 0.7480438351631165
Iter #1716224:  Learning rate = 0.002498:   Batch Loss = 0.980308, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0430864095687866, Accu
racy = 0.7496087551116943
Iter #1720320:  Learning rate = 0.002498:   Batch Loss = 0.956059, Accurac
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.052419662475586, Accur
acy = 0.7313510775566101
Iter #1724416:  Learning rate = 0.002498:   Batch Loss = 1.015545, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0356419086456299, Accu
racy = 0.768909752368927
Iter #1728512:  Learning rate = 0.002498:   Batch Loss = 0.968507, Accurac
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0755890607833862, Accu
racy = 0.738480269908905
Iter #1732608:  Learning rate = 0.002498:   Batch Loss = 1.035042, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.096348524093628, Accur
acy = 0.7621283531188965
Iter #1736704:  Learning rate = 0.002498:   Batch Loss = 1.041259, Accurac
y = 0.76171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0721579790115356, Accu
racy = 0.7473483085632324
Iter #1740800:  Learning rate = 0.002498:   Batch Loss = 1.038128, Accurac
y = 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0565383434295654, Accu
racy = 0.7614327669143677
Iter #1744896:  Learning rate = 0.002498:   Batch Loss = 1.024306, Accurac
y = 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0482994318008423, Accu
racy = 0.7663015127182007
```

```
Iter #1748992:  Learning rate = 0.002498:   Batch Loss = 0.908800, Accurac
y = 0.8125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0207167863845825, Accu
racy = 0.7464788556098938
Iter #1753088:  Learning rate = 0.002498:   Batch Loss = 1.041842, Accurac
y = 0.751953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.076602578163147, Accur
acy = 0.7386541366577148
Iter #1757184:  Learning rate = 0.002498:   Batch Loss = 0.928108, Accurac
y = 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 1.025555968284607, Accur
acy = 0.7713441252708435
Iter #1761280:  Learning rate = 0.002498:   Batch Loss = 0.937429, Accurac
y = 0.80078125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0207650661468506, Accu
racy = 0.751869261264801
Iter #1765376:  Learning rate = 0.002498:   Batch Loss = 0.966179, Accurac
y = 0.779296875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0598074197769165, Accu
racy = 0.7480438351631165
Iter #1769472:  Learning rate = 0.002498:   Batch Loss = 0.918595, Accurac
y = 0.8203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.01669442653656, Accura
cy = 0.778125524520874
Iter #1773568:  Learning rate = 0.002498:   Batch Loss = 0.893030, Accurac
y = 0.830078125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0269782543182373, Accu
racy = 0.7419579029083252
Iter #1777664:  Learning rate = 0.002498:   Batch Loss = 0.963337, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0102384090423584, Accu
racy = 0.7523908615112305
Iter #1781760:  Learning rate = 0.002498:   Batch Loss = 1.017827, Accurac
y = 0.74609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0422606468200684, Accu
racy = 0.758998453617096
Iter #1785856:  Learning rate = 0.002498:   Batch Loss = 0.942958, Accurac
y = 0.810546875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0233336687088013, Accu
racy = 0.7645626664161682
Iter #1789952:  Learning rate = 0.002498:   Batch Loss = 1.044815, Accurac
y = 0.75
PERFORMANCE ON TEST SET:             Batch Loss = 1.041660189628601, Accur
acy = 0.7690836191177368
Iter #1794048:  Learning rate = 0.002498:   Batch Loss = 0.876921, Accurac
y = 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0874338150024414, Accu
racy = 0.7391757965087891
Iter #1798144:  Learning rate = 0.002498:   Batch Loss = 1.171795, Accurac
y = 0.7109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.210585117340088, Accur
acy = 0.6998782753944397
Iter #1802240:  Learning rate = 0.002398:   Batch Loss = 1.024077, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1524709463119507, Accu
racy = 0.718831479549408
Iter #1806336:  Learning rate = 0.002398:   Batch Loss = 0.995012, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0060489177703857, Accu
racy = 0.7558685541152954
Iter #1810432:  Learning rate = 0.002398:   Batch Loss = 1.115390, Accurac
```

```
y = 0.7265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.2471998929977417, Accu
racy = 0.6903147101402283
Iter #1814528:  Learning rate = 0.002398:   Batch Loss = 1.146706, Accurac
y = 0.732421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1741793155670166, Accu
racy = 0.7186576128005981
Iter #1818624:  Learning rate = 0.002398:   Batch Loss = 1.120969, Accurac
y = 0.755859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.128003716468811, Accur
acy = 0.7489132285118103
Iter #1822720:  Learning rate = 0.002398:   Batch Loss = 1.085106, Accurac
y = 0.755859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1459003686904907, Accu
racy = 0.7442184090614319
Iter #1826816:  Learning rate = 0.002398:   Batch Loss = 1.044838, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0701990127563477, Accu
racy = 0.7570857405662537
Iter #1830912:  Learning rate = 0.002398:   Batch Loss = 1.076207, Accurac
y = 0.74609375
PERFORMANCE ON TEST SET:             Batch Loss = 1.086926817893982, Accur
acy = 0.756216287612915
Iter #1835008:  Learning rate = 0.002398:   Batch Loss = 1.082881, Accurac
y = 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1363487243652344, Accu
racy = 0.7419579029083252
Iter #1839104:  Learning rate = 0.002398:   Batch Loss = 1.039790, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0892529487609863, Accu
racy = 0.7461311221122742
Iter #1843200:  Learning rate = 0.002398:   Batch Loss = 1.039903, Accurac
y = 0.751953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.101104497909546, Accur
acy = 0.7492610216140747
Iter #1847296:  Learning rate = 0.002398:   Batch Loss = 1.070918, Accurac
y = 0.73828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1120136976242065, Accu
racy = 0.7410885095596313
Iter #1851392:  Learning rate = 0.002398:   Batch Loss = 1.071955, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0852686166763306, Accu
racy = 0.7544775009155273
Iter #1855488:  Learning rate = 0.002398:   Batch Loss = 1.001025, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0680670738220215, Accu
racy = 0.7593461871147156
Iter #1859584:  Learning rate = 0.002398:   Batch Loss = 1.057543, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.1154470443725586, Accu
racy = 0.7330899238586426
Iter #1863680:  Learning rate = 0.002398:   Batch Loss = 1.101016, Accurac
y = 0.744140625
PERFORMANCE ON TEST SET:             Batch Loss = 1.119731068611145, Accur
acy = 0.7396974563598633
Iter #1867776:  Learning rate = 0.002398:   Batch Loss = 1.016454, Accurac
y = 0.796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0852744579315186, Accu
racy = 0.7525647878646851
Iter #1871872:  Learning rate = 0.002398:   Batch Loss = 1.098576, Accurac
y = 0.740234375
```

```
PERFORMANCE ON TEST SET:            Batch Loss = 1.1819844245910645, Accu
racy = 0.7028343081474304
Iter #1875968:  Learning rate = 0.002398:   Batch Loss = 1.083872, Accurac
y = 0.732421875
PERFORMANCE ON TEST SET:            Batch Loss = 1.0951924324035645, Accu
racy = 0.7576074004173279
Iter #1880064:  Learning rate = 0.002398:   Batch Loss = 1.054193, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:            Batch Loss = 1.1320359706878662, Accu
racy = 0.7536080479621887
Iter #1884160:  Learning rate = 0.002398:   Batch Loss = 1.066842, Accurac
y = 0.751953125
PERFORMANCE ON TEST SET:            Batch Loss = 1.0655856132507324, Accu
racy = 0.7543035745620728
Iter #1888256:  Learning rate = 0.002398:   Batch Loss = 1.063734, Accurac
y = 0.76171875
PERFORMANCE ON TEST SET:            Batch Loss = 1.062270164489746, Accur
acy = 0.7546513676643372
Iter #1892352:  Learning rate = 0.002398:   Batch Loss = 1.204171, Accurac
y = 0.720703125
PERFORMANCE ON TEST SET:            Batch Loss = 1.1487338542938232, Accu
racy = 0.7456094622612
Iter #1896448:  Learning rate = 0.002398:   Batch Loss = 1.005101, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:            Batch Loss = 1.09401273727417, Accura
cy = 0.7543035745620728
Iter #1900544:  Learning rate = 0.002302:   Batch Loss = 1.108179, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:            Batch Loss = 1.0796806812286377, Accu
racy = 0.7546513676643372
Iter #1904640:  Learning rate = 0.002302:   Batch Loss = 1.019745, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:            Batch Loss = 1.0686216354370117, Accu
racy = 0.75969398021698
Iter #1908736:  Learning rate = 0.002302:   Batch Loss = 1.031820, Accurac
y = 0.78125
PERFORMANCE ON TEST SET:            Batch Loss = 1.0527982711791992, Accu
racy = 0.7638671398162842
Iter #1912832:  Learning rate = 0.002302:   Batch Loss = 1.007799, Accurac
y = 0.759765625
PERFORMANCE ON TEST SET:            Batch Loss = 1.0619370937347412, Accu
racy = 0.7683880925178528
Iter #1916928:  Learning rate = 0.002302:   Batch Loss = 1.026513, Accurac
y = 0.751953125
PERFORMANCE ON TEST SET:            Batch Loss = 1.0373209714889526, Accu
racy = 0.747696042060852
Iter #1921024:  Learning rate = 0.002302:   Batch Loss = 1.031970, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:            Batch Loss = 1.0601801872253418, Accu
racy = 0.7612589001655579
Iter #1925120:  Learning rate = 0.002302:   Batch Loss = 1.001532, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:            Batch Loss = 1.0350052118301392, Accu
racy = 0.7602156400680542
Iter #1929216:  Learning rate = 0.002302:   Batch Loss = 1.001947, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:            Batch Loss = 1.0224723815917969, Accu
racy = 0.7643887996673584
Iter #1933312:  Learning rate = 0.002302:   Batch Loss = 1.001350, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:            Batch Loss = 1.023048758506775, Accur
```

```
acy = 0.7683880925178528
Iter #1937408:  Learning rate = 0.002302:   Batch Loss = 0.963398, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0317504405975342, Accu
racy = 0.7623022198677063
Iter #1941504:  Learning rate = 0.002302:   Batch Loss = 0.928064, Accurac
y = 0.8046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0866937637329102, Accu
racy = 0.7334376573562622
Iter #1945600:  Learning rate = 0.002302:   Batch Loss = 1.024252, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0624396800994873, Accu
racy = 0.7216136455535889
Iter #1949696:  Learning rate = 0.002302:   Batch Loss = 0.919386, Accurac
y = 0.80859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0763341188430786, Accu
racy = 0.75969398021698
Iter #1953792:  Learning rate = 0.002302:   Batch Loss = 0.981048, Accurac
y = 0.79296875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0465604066848755, Accu
racy = 0.76334547996521
Iter #1957888:  Learning rate = 0.002302:   Batch Loss = 0.958480, Accurac
y = 0.8046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0269536972045898, Accu
racy = 0.7690836191177368
Iter #1961984:  Learning rate = 0.002302:   Batch Loss = 0.960933, Accurac
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0482852458953857, Accu
racy = 0.7499565482139587
Iter #1966080:  Learning rate = 0.002302:   Batch Loss = 0.927330, Accurac
y = 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0445375442504883, Accu
racy = 0.7536080479621887
Iter #1970176:  Learning rate = 0.002302:   Batch Loss = 0.956164, Accurac
y = 0.8046875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0590386390686035, Accu
racy = 0.7400451898574829
Iter #1974272:  Learning rate = 0.002302:   Batch Loss = 0.972103, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0190339088439941, Accu
racy = 0.7736045718193054
Iter #1978368:  Learning rate = 0.002302:   Batch Loss = 0.884351, Accurac
y = 0.83984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9987333416938782, Accu
racy = 0.7788210511207581
Iter #1982464:  Learning rate = 0.002302:   Batch Loss = 0.964947, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.997748851776123, Accur
acy = 0.7751695513725281
Iter #1986560:  Learning rate = 0.002302:   Batch Loss = 0.963721, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:             Batch Loss = 1.014538049697876, Accur
acy = 0.7508259415626526
Iter #1990656:  Learning rate = 0.002302:   Batch Loss = 1.004386, Accurac
y = 0.76171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0092856884002686, Accu
racy = 0.7501304149627686
Iter #1994752:  Learning rate = 0.002302:   Batch Loss = 0.974055, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.000925064086914, Accur
acy = 0.7732568383216858
```

```
Iter #1998848:  Learning rate = 0.002302:   Batch Loss = 1.000264, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0283187627792358, Accu
racy = 0.7572596073150635
Iter #2002944:  Learning rate = 0.002210:   Batch Loss = 0.973037, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0135406255722046, Accu
racy = 0.7485654950141907
Iter #2007040:  Learning rate = 0.002210:   Batch Loss = 0.990888, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0437569618225098, Accu
racy = 0.7565640807151794
Iter #2011136:  Learning rate = 0.002210:   Batch Loss = 1.026631, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0918889045715332, Accu
racy = 0.7593461871147156
Iter #2015232:  Learning rate = 0.002210:   Batch Loss = 1.005755, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0359697341918945, Accu
racy = 0.7652581930160522
Iter #2019328:  Learning rate = 0.002210:   Batch Loss = 0.926965, Accurac
y = 0.80859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0463606119155884, Accu
racy = 0.754825234413147
Iter #2023424:  Learning rate = 0.002210:   Batch Loss = 0.984875, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.048729658126831, Accur
acy = 0.7426534295082092
Iter #2027520:  Learning rate = 0.002210:   Batch Loss = 0.941919, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9894427061080933, Accu
racy = 0.7826464772224426
Iter #2031616:  Learning rate = 0.002210:   Batch Loss = 1.011278, Accurac
y = 0.748046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9961905479431152, Accu
racy = 0.7471743822097778
Iter #2035712:  Learning rate = 0.002210:   Batch Loss = 0.982830, Accurac
y = 0.78515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9970383644104004, Accu
racy = 0.7703008055686951
Iter #2039808:  Learning rate = 0.002210:   Batch Loss = 1.047994, Accurac
y = 0.744140625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1281318664550781, Accu
racy = 0.7264823317527771
Iter #2043904:  Learning rate = 0.002210:   Batch Loss = 1.049932, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0017271041870117, Accu
racy = 0.7793427109718323
Iter #2048000:  Learning rate = 0.002210:   Batch Loss = 1.068969, Accurac
y = 0.75390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0178475379943848, Accu
racy = 0.7623022198677063
Iter #2052096:  Learning rate = 0.002210:   Batch Loss = 0.968532, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0511412620544434, Accu
racy = 0.7703008055686951
Iter #2056192:  Learning rate = 0.002210:   Batch Loss = 0.943798, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0545878410339355, Accu
racy = 0.7532603144645691
Iter #2060288:  Learning rate = 0.002210:   Batch Loss = 0.971445, Accurac
```

```
y = 0.79296875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0304360389709473, Accu
racy = 0.7715179920196533
Iter #2064384:  Learning rate = 0.002210:   Batch Loss = 0.977505, Accurac
y = 0.79296875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0180691480636597, Accu
racy = 0.7687358856201172
Iter #2068480:  Learning rate = 0.002210:   Batch Loss = 1.006278, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9955016374588013, Accu
racy = 0.7676925659179688
Iter #2072576:  Learning rate = 0.002210:   Batch Loss = 0.954452, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.021625280380249, Accur
acy = 0.7706485986709595
Iter #2076672:  Learning rate = 0.002210:   Batch Loss = 0.959162, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 1.1162168979644775, Accu
racy = 0.7350026369094849
Iter #2080768:  Learning rate = 0.002210:   Batch Loss = 1.038103, Accurac
y = 0.755859375
PERFORMANCE ON TEST SET:             Batch Loss = 1.034006118774414, Accur
acy = 0.7623022198677063
Iter #2084864:  Learning rate = 0.002210:   Batch Loss = 0.961444, Accurac
y = 0.78515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.077413558959961, Accur
acy = 0.7287428379058838
Iter #2088960:  Learning rate = 0.002210:   Batch Loss = 0.958494, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.025784969329834, Accur
acy = 0.77742999792099
Iter #2093056:  Learning rate = 0.002210:   Batch Loss = 0.900897, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0090522766113281, Accu
racy = 0.7657798528671265
Iter #2097152:  Learning rate = 0.002210:   Batch Loss = 0.974839, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0482349395751953, Accu
racy = 0.7654321193695068
Iter #2101248:  Learning rate = 0.002122:   Batch Loss = 0.945166, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 1.015492558479309, Accur
acy = 0.747696042060852
Iter #2105344:  Learning rate = 0.002122:   Batch Loss = 0.979382, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0508365631103516, Accu
racy = 0.7483915686607361
Iter #2109440:  Learning rate = 0.002122:   Batch Loss = 0.951770, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0185986757278442, Accu
racy = 0.7520431280136108
Iter #2113536:  Learning rate = 0.002122:   Batch Loss = 0.942012, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0178945064544678, Accu
racy = 0.7758650779724121
Iter #2117632:  Learning rate = 0.002122:   Batch Loss = 0.931853, Accurac
y = 0.79296875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0340033769607544, Accu
racy = 0.7722135186195374
Iter #2121728:  Learning rate = 0.002122:   Batch Loss = 0.947470, Accurac
y = 0.765625
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.9888512492179871, Accu
racy = 0.7546513676643372
Iter #2125824:  Learning rate = 0.002122:   Batch Loss = 1.046025, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0450273752212524, Accu
racy = 0.7723873853683472
Iter #2129920:  Learning rate = 0.002122:   Batch Loss = 0.974577, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0137276649475098, Accu
racy = 0.7663015127182007
Iter #2134016:  Learning rate = 0.002122:   Batch Loss = 0.918893, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.028969645500183, Accur
acy = 0.7482177019119263
Iter #2138112:  Learning rate = 0.002122:   Batch Loss = 0.993924, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9956344366073608, Accu
racy = 0.7763867378234863
Iter #2142208:  Learning rate = 0.002122:   Batch Loss = 0.976834, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9963197708129883, Accu
racy = 0.7609111666679382
Iter #2146304:  Learning rate = 0.002122:   Batch Loss = 1.022440, Accurac
y = 0.759765625
PERFORMANCE ON TEST SET:             Batch Loss = 1.077785611152649, Accur
acy = 0.7461311221122742
Iter #2150400:  Learning rate = 0.002122:   Batch Loss = 1.018135, Accurac
y = 0.76171875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1057053804397583, Accu
racy = 0.733611524105072
Iter #2154496:  Learning rate = 0.002122:   Batch Loss = 0.936616, Accurac
y = 0.796875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0274403095245361, Accu
racy = 0.746304988861084
Iter #2158592:  Learning rate = 0.002122:   Batch Loss = 1.005494, Accurac
y = 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 1.03566312789917, Accura
cy = 0.7478699088096619
Iter #2162688:  Learning rate = 0.002122:   Batch Loss = 0.981519, Accurac
y = 0.779296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9871788620948792, Accu
racy = 0.7819509506225586
Iter #2166784:  Learning rate = 0.002122:   Batch Loss = 1.012414, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9877625107765198, Accu
racy = 0.773082971572876
Iter #2170880:  Learning rate = 0.002122:   Batch Loss = 1.053598, Accurac
y = 0.72265625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0121536254882812, Accu
racy = 0.7673448324203491
Iter #2174976:  Learning rate = 0.002122:   Batch Loss = 0.974429, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0765548944473267, Accu
racy = 0.7551730275154114
Iter #2179072:  Learning rate = 0.002122:   Batch Loss = 0.965828, Accurac
y = 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9925293922424316, Accu
racy = 0.7668231725692749
Iter #2183168:  Learning rate = 0.002122:   Batch Loss = 0.970210, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0121475458145142, Accu
```

```
racy = 0.76334547996521
Iter #2187264:  Learning rate = 0.002122:   Batch Loss = 0.986086, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0032482147216797, Accu
racy = 0.7739523649215698
Iter #2191360:  Learning rate = 0.002122:   Batch Loss = 0.907663, Accurac
y = 0.81640625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0080909729003906, Accu
racy = 0.774474024772644
Iter #2195456:  Learning rate = 0.002122:   Batch Loss = 0.917918, Accurac
y = 0.796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9971964359283447, Accu
racy = 0.7833420038223267
Iter #2199552:  Learning rate = 0.002122:   Batch Loss = 0.987444, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9875388145446777, Accu
racy = 0.7723873853683472
Iter #2203648:  Learning rate = 0.002037:   Batch Loss = 0.938364, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9907670021057129, Accu
racy = 0.7678664326667786
Iter #2207744:  Learning rate = 0.002037:   Batch Loss = 0.937403, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9733511805534363, Accu
racy = 0.773082971572876
Iter #2211840:  Learning rate = 0.002037:   Batch Loss = 0.965919, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9606539011001587, Accu
racy = 0.7840375304222107
Iter #2215936:  Learning rate = 0.002037:   Batch Loss = 0.937528, Accurac
y = 0.810546875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0493288040161133, Accu
racy = 0.7623022198677063
Iter #2220032:  Learning rate = 0.002037:   Batch Loss = 0.931129, Accurac
y = 0.80078125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0551106929779053, Accu
racy = 0.7483915686607361
Iter #2224128:  Learning rate = 0.002037:   Batch Loss = 1.014255, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0259565114974976, Accu
racy = 0.7649104595184326
Iter #2228224:  Learning rate = 0.002037:   Batch Loss = 0.967787, Accurac
y = 0.779296875
PERFORMANCE ON TEST SET:             Batch Loss = 1.045028805732727, Accur
acy = 0.7449139356613159
Iter #2232320:  Learning rate = 0.002037:   Batch Loss = 0.956393, Accurac
y = 0.78515625
PERFORMANCE ON TEST SET:             Batch Loss = 1.003172755241394, Accur
acy = 0.7706485986709595
Iter #2236416:  Learning rate = 0.002037:   Batch Loss = 0.925455, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9952754974365234, Accu
racy = 0.7699530720710754
Iter #2240512:  Learning rate = 0.002037:   Batch Loss = 1.056256, Accurac
y = 0.75
PERFORMANCE ON TEST SET:             Batch Loss = 1.102189540863037, Accur
acy = 0.7546513676643372
Iter #2244608:  Learning rate = 0.002037:   Batch Loss = 0.968221, Accurac
y = 0.76953125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0157208442687988, Accu
racy = 0.7541297078132629
```

```
Iter #2248704:  Learning rate = 0.002037:   Batch Loss = 0.939826, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9877668619155884, Accu
racy = 0.7715179920196533
Iter #2252800:  Learning rate = 0.002037:   Batch Loss = 0.907282, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9757534861564636, Accu
racy = 0.7605633735656738
Iter #2256896:  Learning rate = 0.002037:   Batch Loss = 0.925465, Accurac
y = 0.814453125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0063116550445557, Accu
racy = 0.7584767937660217
Iter #2260992:  Learning rate = 0.002037:   Batch Loss = 0.931171, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9761041402816772, Accu
racy = 0.7796905040740967
Iter #2265088:  Learning rate = 0.002037:   Batch Loss = 0.915680, Accurac
y = 0.80078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.975098729133606, Accur
acy = 0.7715179920196533
Iter #2269184:  Learning rate = 0.002037:   Batch Loss = 0.914232, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9902149438858032, Accu
racy = 0.7765606045722961
Iter #2273280:  Learning rate = 0.002037:   Batch Loss = 0.933486, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9947229623794556, Accu
racy = 0.7443922758102417
Iter #2277376:  Learning rate = 0.002037:   Batch Loss = 0.973277, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9648227095603943, Accu
racy = 0.782994270324707
Iter #2281472:  Learning rate = 0.002037:   Batch Loss = 0.954866, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9632856249809265, Accu
racy = 0.7800382375717163
Iter #2285568:  Learning rate = 0.002037:   Batch Loss = 0.929365, Accurac
y = 0.78125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9971973896026611, Accu
racy = 0.7513476014137268
Iter #2289664:  Learning rate = 0.002037:   Batch Loss = 0.942944, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9862554669380188, Accu
racy = 0.7628238797187805
Iter #2293760:  Learning rate = 0.002037:   Batch Loss = 0.850651, Accurac
y = 0.8046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9996663331985474, Accu
racy = 0.7607372403144836
Iter #2297856:  Learning rate = 0.002037:   Batch Loss = 1.001151, Accurac
y = 0.767578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9806104302406311, Accu
racy = 0.7765606045722961
Iter #2301952:  Learning rate = 0.001955:   Batch Loss = 0.887576, Accurac
y = 0.828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9820981025695801, Accu
racy = 0.7685620188713074
Iter #2306048:  Learning rate = 0.001955:   Batch Loss = 0.967450, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9810292720794678, Accu
racy = 0.7749956250190735
Iter #2310144:  Learning rate = 0.001955:   Batch Loss = 0.914782, Accurac
```

```
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9925390481948853, Accu
racy = 0.7716918587684631
Iter #2314240:  Learning rate = 0.001955:   Batch Loss = 0.981508, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9802835583686829, Accu
racy = 0.747696042060852
Iter #2318336:  Learning rate = 0.001955:   Batch Loss = 0.880130, Accurac
y = 0.830078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9523251056671143, Accu
racy = 0.7819509506225586
Iter #2322432:  Learning rate = 0.001955:   Batch Loss = 0.890260, Accurac
y = 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9778720736503601, Accu
racy = 0.7584767937660217
Iter #2326528:  Learning rate = 0.001955:   Batch Loss = 0.950215, Accurac
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9842774868011475, Accu
racy = 0.7619544267654419
Iter #2330624:  Learning rate = 0.001955:   Batch Loss = 0.936448, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9676327705383301, Accu
racy = 0.7629977464675903
Iter #2334720:  Learning rate = 0.001955:   Batch Loss = 0.910583, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.012549877166748, Accur
acy = 0.7709963321685791
Iter #2338816:  Learning rate = 0.001955:   Batch Loss = 0.876138, Accurac
y = 0.822265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9851580858230591, Accu
racy = 0.7729090452194214
Iter #2342912:  Learning rate = 0.001955:   Batch Loss = 0.972950, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9993798732757568, Accu
racy = 0.7708224654197693
Iter #2347008:  Learning rate = 0.001955:   Batch Loss = 0.944065, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0347352027893066, Accu
racy = 0.7490870952606201
Iter #2351104:  Learning rate = 0.001955:   Batch Loss = 0.974027, Accurac
y = 0.76171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9999964833259583, Accu
racy = 0.76334547996521
Iter #2355200:  Learning rate = 0.001955:   Batch Loss = 0.902958, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9532148241996765, Accu
racy = 0.7784733176231384
Iter #2359296:  Learning rate = 0.001955:   Batch Loss = 0.971758, Accurac
y = 0.77734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.0003399848937988, Accu
racy = 0.774474024772644
Iter #2363392:  Learning rate = 0.001955:   Batch Loss = 0.914678, Accurac
y = 0.78125
PERFORMANCE ON TEST SET:             Batch Loss = 0.96930330991745, Accura
cy = 0.7711702585220337
Iter #2367488:  Learning rate = 0.001955:   Batch Loss = 0.913509, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9734935760498047, Accu
racy = 0.7544775009155273
Iter #2371584:  Learning rate = 0.001955:   Batch Loss = 0.949834, Accurac
y = 0.78515625
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.9881112575531006, Accu
racy = 0.7729090452194214
Iter #2375680:  Learning rate = 0.001955:   Batch Loss = 0.883131, Accurac
y = 0.796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9603437185287476, Accu
racy = 0.7732568383216858
Iter #2379776:  Learning rate = 0.001955:   Batch Loss = 0.909723, Accurac
y = 0.79296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9815105199813843, Accu
racy = 0.7765606045722961
Iter #2383872:  Learning rate = 0.001955:   Batch Loss = 0.849972, Accurac
y = 0.826171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.972668468952179, Accur
acy = 0.7777777910232544
Iter #2387968:  Learning rate = 0.001955:   Batch Loss = 0.901400, Accurac
y = 0.8046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9781492948532104, Accu
racy = 0.7729090452194214
Iter #2392064:  Learning rate = 0.001955:   Batch Loss = 0.909856, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9773904085159302, Accu
racy = 0.7739523649215698
Iter #2396160:  Learning rate = 0.001955:   Batch Loss = 0.889710, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9799371957778931, Accu
racy = 0.7812554240226746
Iter #2400256:  Learning rate = 0.001877:   Batch Loss = 0.909652, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9715712070465088, Accu
racy = 0.7796905040740967
Iter #2404352:  Learning rate = 0.001877:   Batch Loss = 0.932570, Accurac
y = 0.80859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9794715642929077, Accu
racy = 0.7720396518707275
Iter #2408448:  Learning rate = 0.001877:   Batch Loss = 0.932275, Accurac
y = 0.779296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9587781429290771, Accu
racy = 0.7814292907714844
Iter #2412544:  Learning rate = 0.001877:   Batch Loss = 0.838300, Accurac
y = 0.806640625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0043649673461914, Accu
racy = 0.7708224654197693
Iter #2416640:  Learning rate = 0.001877:   Batch Loss = 0.913777, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9584262371063232, Accu
racy = 0.7798643708229065
Iter #2420736:  Learning rate = 0.001877:   Batch Loss = 0.935351, Accurac
y = 0.78515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.957136869430542, Accur
acy = 0.7810815572738647
Iter #2424832:  Learning rate = 0.001877:   Batch Loss = 0.905191, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9625738859176636, Accu
racy = 0.7854286432266235
Iter #2428928:  Learning rate = 0.001877:   Batch Loss = 0.979895, Accurac
y = 0.755859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9645591974258423, Accu
racy = 0.7570857405662537
Iter #2433024:  Learning rate = 0.001877:   Batch Loss = 0.895255, Accurac
y = 0.79296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9692785739898682, Accu
```

```
racy = 0.7758650779724121
Iter #2437120:  Learning rate = 0.001877:   Batch Loss = 0.908655, Accurac
y = 0.80078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9815523624420166, Accu
racy = 0.7614327669143677
Iter #2441216:  Learning rate = 0.001877:   Batch Loss = 0.907403, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9521104097366333, Accu
racy = 0.7786471843719482
Iter #2445312:  Learning rate = 0.001877:   Batch Loss = 0.907049, Accurac
y = 0.806640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.977770209312439, Accur
acy = 0.7777777910232544
Iter #2449408:  Learning rate = 0.001877:   Batch Loss = 0.896258, Accurac
y = 0.78515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.978635847568512, Accur
acy = 0.77377849817276
Iter #2453504:  Learning rate = 0.001877:   Batch Loss = 0.912642, Accurac
y = 0.8046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9845306873321533, Accu
racy = 0.7796905040740967
Iter #2457600:  Learning rate = 0.001877:   Batch Loss = 0.960504, Accurac
y = 0.7734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9736341834068298, Accu
racy = 0.7720396518707275
Iter #2461696:  Learning rate = 0.001877:   Batch Loss = 0.935981, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9641077518463135, Accu
racy = 0.7849069833755493
Iter #2465792:  Learning rate = 0.001877:   Batch Loss = 0.903369, Accurac
y = 0.79296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9833476543426514, Accu
racy = 0.7758650779724121
Iter #2469888:  Learning rate = 0.001877:   Batch Loss = 0.991417, Accurac
y = 0.75
PERFORMANCE ON TEST SET:             Batch Loss = 1.0793087482452393, Accu
racy = 0.7181359529495239
Iter #2473984:  Learning rate = 0.001877:   Batch Loss = 0.979665, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 1.036872386932373, Accur
acy = 0.76334547996521
Iter #2478080:  Learning rate = 0.001877:   Batch Loss = 0.962062, Accurac
y = 0.75390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9698967337608337, Accu
racy = 0.7661276459693909
Iter #2482176:  Learning rate = 0.001877:   Batch Loss = 0.907971, Accurac
y = 0.78515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9616332650184631, Accu
racy = 0.773082971572876
Iter #2486272:  Learning rate = 0.001877:   Batch Loss = 0.909697, Accurac
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9638713002204895, Accu
racy = 0.7609111666679382
Iter #2490368:  Learning rate = 0.001877:   Batch Loss = 0.885551, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9811505079269409, Accu
racy = 0.7782994508743286
Iter #2494464:  Learning rate = 0.001877:   Batch Loss = 0.871301, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 1.000397801399231, Accur
acy = 0.7683880925178528
```

```
Iter #2498560:  Learning rate = 0.001877:   Batch Loss = 0.968552, Accurac
y = 0.765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9906219840049744, Accu
racy = 0.7574334740638733
Iter #2502656:  Learning rate = 0.001802:   Batch Loss = 0.871149, Accurac
y = 0.80078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9696425199508667, Accu
racy = 0.7569118142127991
Iter #2506752:  Learning rate = 0.001802:   Batch Loss = 0.994823, Accurac
y = 0.751953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9508850574493408, Accu
racy = 0.7788210511207581
Iter #2510848:  Learning rate = 0.001802:   Batch Loss = 0.874952, Accurac
y = 0.80859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9781609773635864, Accu
racy = 0.769605278968811
Iter #2514944:  Learning rate = 0.001802:   Batch Loss = 0.852007, Accurac
y = 0.826171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9693976640701294, Accu
racy = 0.7798643708229065
Iter #2519040:  Learning rate = 0.001802:   Batch Loss = 0.895773, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9647326469421387, Accu
racy = 0.7718657851219177
Iter #2523136:  Learning rate = 0.001802:   Batch Loss = 0.875610, Accurac
y = 0.826171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9535009860992432, Accu
racy = 0.7782994508743286
Iter #2527232:  Learning rate = 0.001802:   Batch Loss = 0.956226, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.937940239906311, Accur
acy = 0.7871674299240112
Iter #2531328:  Learning rate = 0.001802:   Batch Loss = 0.881930, Accurac
y = 0.8046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9421074986457825, Accu
racy = 0.7835159301757812
Iter #2535424:  Learning rate = 0.001802:   Batch Loss = 1.077370, Accurac
y = 0.732421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0800448656082153, Accu
racy = 0.7527386546134949
Iter #2539520:  Learning rate = 0.001802:   Batch Loss = 0.911880, Accurac
y = 0.7890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9547363519668579, Accu
racy = 0.7619544267654419
Iter #2543616:  Learning rate = 0.001802:   Batch Loss = 0.933781, Accurac
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.0478134155273438, Accu
racy = 0.7438706159591675
Iter #2547712:  Learning rate = 0.001802:   Batch Loss = 1.263568, Accurac
y = 0.681640625
PERFORMANCE ON TEST SET:             Batch Loss = 1.0665887594223022, Accu
racy = 0.7543035745620728
Iter #2551808:  Learning rate = 0.001802:   Batch Loss = 0.870874, Accurac
y = 0.8046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9980449676513672, Accu
racy = 0.7499565482139587
Iter #2555904:  Learning rate = 0.001802:   Batch Loss = 0.900509, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9492555856704712, Accu
racy = 0.7793427109718323
Iter #2560000:  Learning rate = 0.001802:   Batch Loss = 0.842908, Accurac
```

```
y = 0.8359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9587193727493286, Accu
racy = 0.7803860306739807
Iter #2564096:  Learning rate = 0.001802:   Batch Loss = 0.909572, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9559444785118103, Accu
racy = 0.7798643708229065
Iter #2568192:  Learning rate = 0.001802:   Batch Loss = 0.916295, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9429531097412109, Accu
racy = 0.7711702585220337
Iter #2572288:  Learning rate = 0.001802:   Batch Loss = 0.954699, Accurac
y = 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9701017141342163, Accu
racy = 0.7795165777206421
Iter #2576384:  Learning rate = 0.001802:   Batch Loss = 0.916667, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9444547891616821, Accu
racy = 0.7645626664161682
Iter #2580480:  Learning rate = 0.001802:   Batch Loss = 0.921111, Accurac
y = 0.78125
PERFORMANCE ON TEST SET:             Batch Loss = 0.943974494934082, Accur
acy = 0.76334547996521
Iter #2584576:  Learning rate = 0.001802:   Batch Loss = 0.891846, Accurac
y = 0.78125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9482229948043823, Accu
racy = 0.7810815572738647
Iter #2588672:  Learning rate = 0.001802:   Batch Loss = 0.888904, Accurac
y = 0.810546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9546190500259399, Accu
racy = 0.7856025099754333
Iter #2592768:  Learning rate = 0.001802:   Batch Loss = 0.889956, Accurac
y = 0.791015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9343320727348328, Accu
racy = 0.7845591902732849
Iter #2596864:  Learning rate = 0.001802:   Batch Loss = 0.891290, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 1.1295678615570068, Accu
racy = 0.7318727374076843
Iter #2600960:  Learning rate = 0.001730:   Batch Loss = 0.904026, Accurac
y = 0.779296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9664934873580933, Accu
racy = 0.7442184090614319
Iter #2605056:  Learning rate = 0.001730:   Batch Loss = 0.866885, Accurac
y = 0.828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9530532360076904, Accu
racy = 0.7746478915214539
Iter #2609152:  Learning rate = 0.001730:   Batch Loss = 0.884870, Accurac
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9407984018325806, Accu
racy = 0.77377849817276
Iter #2613248:  Learning rate = 0.001730:   Batch Loss = 0.949248, Accurac
y = 0.775390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.9415466785430908, Accu
racy = 0.7849069833755493
Iter #2617344:  Learning rate = 0.001730:   Batch Loss = 0.906347, Accurac
y = 0.779296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9152872562408447, Accu
racy = 0.7809076905250549
Iter #2621440:  Learning rate = 0.001730:   Batch Loss = 0.858598, Accurac
y = 0.80078125
```

```
PERFORMANCE ON TEST SET:            Batch Loss = 0.9369837045669556, Accu
racy = 0.7878629565238953
Iter #2625536:  Learning rate = 0.001730:   Batch Loss = 0.875881, Accurac
y = 0.830078125
PERFORMANCE ON TEST SET:            Batch Loss = 0.9564520120620728, Accu
racy = 0.7699530720710754
Iter #2629632:  Learning rate = 0.001730:   Batch Loss = 0.823438, Accurac
y = 0.8203125
PERFORMANCE ON TEST SET:            Batch Loss = 0.8802493810653687, Accu
racy = 0.8002086877822876
Iter #2633728:  Learning rate = 0.001730:   Batch Loss = 0.903128, Accurac
y = 0.787109375
PERFORMANCE ON TEST SET:            Batch Loss = 0.9456185102462769, Accu
racy = 0.75969398021698
Iter #2637824:  Learning rate = 0.001730:   Batch Loss = 0.821291, Accurac
y = 0.82421875
PERFORMANCE ON TEST SET:            Batch Loss = 0.8861021399497986, Accu
racy = 0.7845591902732849
Iter #2641920:  Learning rate = 0.001730:   Batch Loss = 0.841907, Accurac
y = 0.82421875
PERFORMANCE ON TEST SET:            Batch Loss = 0.8954052925109863, Accu
racy = 0.7864719033241272
Iter #2646016:  Learning rate = 0.001730:   Batch Loss = 0.836296, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:            Batch Loss = 0.8661041855812073, Accu
racy = 0.8059467673301697
Iter #2650112:  Learning rate = 0.001730:   Batch Loss = 0.826849, Accurac
y = 0.828125
PERFORMANCE ON TEST SET:            Batch Loss = 0.8939279317855835, Accu
racy = 0.7930794358253479
Iter #2654208:  Learning rate = 0.001730:   Batch Loss = 0.889475, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:            Batch Loss = 0.9310283064842224, Accu
racy = 0.7650843262672424
Iter #2658304:  Learning rate = 0.001730:   Batch Loss = 0.843154, Accurac
y = 0.814453125
PERFORMANCE ON TEST SET:            Batch Loss = 0.8871028423309326, Accu
racy = 0.7998608946800232
Iter #2662400:  Learning rate = 0.001730:   Batch Loss = 0.859468, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:            Batch Loss = 0.8735517263412476, Accu
racy = 0.7889062762260437
Iter #2666496:  Learning rate = 0.001730:   Batch Loss = 0.853007, Accurac
y = 0.82421875
PERFORMANCE ON TEST SET:            Batch Loss = 0.892049252986908, Accur
acy = 0.7786471843719482
Iter #2670592:  Learning rate = 0.001730:   Batch Loss = 0.873820, Accurac
y = 0.806640625
PERFORMANCE ON TEST SET:            Batch Loss = 0.893176794052124, Accur
acy = 0.7725613117218018
Iter #2674688:  Learning rate = 0.001730:   Batch Loss = 0.799087, Accurac
y = 0.8203125
PERFORMANCE ON TEST SET:            Batch Loss = 0.8748794794082642, Accu
racy = 0.8045557141304016
Iter #2678784:  Learning rate = 0.001730:   Batch Loss = 0.841008, Accurac
y = 0.833984375
PERFORMANCE ON TEST SET:            Batch Loss = 0.9053493142127991, Accu
racy = 0.7974265217781067
Iter #2682880:  Learning rate = 0.001730:   Batch Loss = 0.739572, Accurac
y = 0.865234375
PERFORMANCE ON TEST SET:            Batch Loss = 0.8997687697410583, Accu
```

```
racy = 0.8017736077308655
Iter #2686976:  Learning rate = 0.001730:   Batch Loss = 0.817024, Accurac
y = 0.814453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.8873555064201355, Accu
racy = 0.7763867378234863
Iter #2691072:  Learning rate = 0.001730:   Batch Loss = 0.824966, Accurac
y = 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8685689568519592, Accu
racy = 0.8094244599342346
Iter #2695168:  Learning rate = 0.001730:   Batch Loss = 0.805101, Accurac
y = 0.853515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.8506677150726318, Accu
racy = 0.8167275190353394
Iter #2699264:  Learning rate = 0.001730:   Batch Loss = 0.803483, Accurac
y = 0.8359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8428695201873779, Accu
racy = 0.8111632466316223
Iter #2703360:  Learning rate = 0.001661:   Batch Loss = 0.755635, Accurac
y = 0.8515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.862490177154541, Accur
acy = 0.8118588328361511
Iter #2707456:  Learning rate = 0.001661:   Batch Loss = 0.779442, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8383185863494873, Accu
racy = 0.818988025188446
Iter #2711552:  Learning rate = 0.001661:   Batch Loss = 0.795315, Accurac
y = 0.8203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.8688585758209229, Accu
racy = 0.8090766668319702
Iter #2715648:  Learning rate = 0.001661:   Batch Loss = 0.749914, Accurac
y = 0.853515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.843782901763916, Accur
acy = 0.8076856136322021
Iter #2719744:  Learning rate = 0.001661:   Batch Loss = 0.784643, Accurac
y = 0.833984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8733795285224915, Accu
racy = 0.8153364658355713
Iter #2723840:  Learning rate = 0.001661:   Batch Loss = 0.781164, Accurac
y = 0.849609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.9426486492156982, Accu
racy = 0.7988175749778748
Iter #2727936:  Learning rate = 0.001661:   Batch Loss = 0.796433, Accurac
y = 0.81640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.8472400903701782, Accu
racy = 0.8186402320861816
Iter #2732032:  Learning rate = 0.001661:   Batch Loss = 0.737072, Accurac
y = 0.85546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.8431919813156128, Accu
racy = 0.8228134512901306
Iter #2736128:  Learning rate = 0.001661:   Batch Loss = 0.777000, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8360099792480469, Accu
racy = 0.8210746049880981
Iter #2740224:  Learning rate = 0.001661:   Batch Loss = 0.785217, Accurac
y = 0.82421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.8579179048538208, Accu
racy = 0.80559903383255
Iter #2744320:  Learning rate = 0.001661:   Batch Loss = 0.796419, Accurac
y = 0.822265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.8171462416648865, Accu
racy = 0.8266388177871704
```

```
Iter #2748416:  Learning rate = 0.001661:   Batch Loss = 0.781689, Accurac
y = 0.830078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.8742567300796509, Accu
racy = 0.8080334067344666
Iter #2752512:  Learning rate = 0.001661:   Batch Loss = 0.771094, Accurac
y = 0.86328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.8581210374832153, Accu
racy = 0.7986437082290649
Iter #2756608:  Learning rate = 0.001661:   Batch Loss = 0.704043, Accurac
y = 0.880859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8371187448501587, Accu
racy = 0.8087289333343506
Iter #2760704:  Learning rate = 0.001661:   Batch Loss = 0.711408, Accurac
y = 0.873046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.8073700666427612, Accu
racy = 0.8313336968421936
Iter #2764800:  Learning rate = 0.001661:   Batch Loss = 0.730238, Accurac
y = 0.828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7861325740814209, Accu
racy = 0.8412449955940247
Iter #2768896:  Learning rate = 0.001661:   Batch Loss = 0.694603, Accurac
y = 0.859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7753654718399048, Accu
racy = 0.8447226285934448
Iter #2772992:  Learning rate = 0.001661:   Batch Loss = 0.735093, Accurac
y = 0.859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.795419454574585, Accur
acy = 0.8388106226921082
Iter #2777088:  Learning rate = 0.001661:   Batch Loss = 0.733902, Accurac
y = 0.859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8087548017501831, Accu
racy = 0.8254216909408569
Iter #2781184:  Learning rate = 0.001661:   Batch Loss = 0.759031, Accurac
y = 0.86328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7879842519760132, Accu
racy = 0.8292471170425415
Iter #2785280:  Learning rate = 0.001661:   Batch Loss = 0.695240, Accurac
y = 0.859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7878859043121338, Accu
racy = 0.8428099751472473
Iter #2789376:  Learning rate = 0.001661:   Batch Loss = 0.738152, Accurac
y = 0.857421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7698531150817871, Accu
racy = 0.8365501761436462
Iter #2793472:  Learning rate = 0.001661:   Batch Loss = 0.713278, Accurac
y = 0.87109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8005120754241943, Accu
racy = 0.8309859037399292
Iter #2797568:  Learning rate = 0.001661:   Batch Loss = 0.669721, Accurac
y = 0.88671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7585604190826416, Accu
racy = 0.8624587059020996
Iter #2801664:  Learning rate = 0.001594:   Batch Loss = 0.756000, Accurac
y = 0.845703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7524011135101318, Accu
racy = 0.8450704216957092
Iter #2805760:  Learning rate = 0.001594:   Batch Loss = 0.737781, Accurac
y = 0.8671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.849351167678833, Accur
acy = 0.8255955576896667
Iter #2809856:  Learning rate = 0.001594:   Batch Loss = 0.733385, Accurac
```

```
y = 0.869140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7642515897750854, Accu
racy = 0.8398539423942566
Iter #2813952:  Learning rate = 0.001594:   Batch Loss = 0.743442, Accurac
y = 0.86328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7638765573501587, Accu
racy = 0.8391584157943726
Iter #2818048:  Learning rate = 0.001594:   Batch Loss = 0.751699, Accurac
y = 0.849609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7553064823150635, Accu
racy = 0.8417666554450989
Iter #2822144:  Learning rate = 0.001594:   Batch Loss = 0.686847, Accurac
y = 0.8828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7290030717849731, Accu
racy = 0.86576247215271
Iter #2826240:  Learning rate = 0.001594:   Batch Loss = 0.684098, Accurac
y = 0.8828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.760366678237915, Accur
acy = 0.8471570014953613
Iter #2830336:  Learning rate = 0.001594:   Batch Loss = 0.675206, Accurac
y = 0.8828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7249462604522705, Accu
racy = 0.8676751852035522
Iter #2834432:  Learning rate = 0.001594:   Batch Loss = 0.667653, Accurac
y = 0.87890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7261780500411987, Accu
racy = 0.8697617650032043
Iter #2838528:  Learning rate = 0.001594:   Batch Loss = 0.650376, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7170584201812744, Accu
racy = 0.8741088509559631
Iter #2842624:  Learning rate = 0.001594:   Batch Loss = 0.731588, Accurac
y = 0.857421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.724698007106781, Accur
acy = 0.8603721261024475
Iter #2846720:  Learning rate = 0.001594:   Batch Loss = 0.657471, Accurac
y = 0.865234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7193485498428345, Accu
racy = 0.870631217956543
Iter #2850816:  Learning rate = 0.001594:   Batch Loss = 0.664121, Accurac
y = 0.88671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7240687012672424, Accu
racy = 0.871326744556427
Iter #2854912:  Learning rate = 0.001594:   Batch Loss = 0.732218, Accurac
y = 0.857421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7437025308609009, Accu
racy = 0.8638497591018677
Iter #2859008:  Learning rate = 0.001594:   Batch Loss = 0.648815, Accurac
y = 0.888671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.750897228717804, Accur
acy = 0.8648930788040161
Iter #2863104:  Learning rate = 0.001594:   Batch Loss = 0.658882, Accurac
y = 0.8828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.726648211479187, Accur
acy = 0.8647191524505615
Iter #2867200:  Learning rate = 0.001594:   Batch Loss = 0.653445, Accurac
y = 0.89453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7521296739578247, Accu
racy = 0.8485480546951294
Iter #2871296:  Learning rate = 0.001594:   Batch Loss = 0.651945, Accurac
y = 0.89453125
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.7381912469863892, Accu
racy = 0.8513302206993103
Iter #2875392:  Learning rate = 0.001594:   Batch Loss = 0.624962, Accurac
y = 0.8984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7016326189041138, Accu
racy = 0.8810641765594482
Iter #2879488:  Learning rate = 0.001594:   Batch Loss = 0.697106, Accurac
y = 0.873046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7222169041633606, Accu
racy = 0.8699356913566589
Iter #2883584:  Learning rate = 0.001594:   Batch Loss = 0.668588, Accurac
y = 0.88671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7179802656173706, Accu
racy = 0.871326744556427
Iter #2887680:  Learning rate = 0.001594:   Batch Loss = 0.746529, Accurac
y = 0.861328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7486504316329956, Accu
racy = 0.8619370460510254
Iter #2891776:  Learning rate = 0.001594:   Batch Loss = 0.675991, Accurac
y = 0.876953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7167770266532898, Accu
racy = 0.8761954307556152
Iter #2895872:  Learning rate = 0.001594:   Batch Loss = 0.671471, Accurac
y = 0.87890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.746688723564148, Accur
acy = 0.8589810729026794
Iter #2899968:  Learning rate = 0.001594:   Batch Loss = 0.684603, Accurac
y = 0.873046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7378972768783569, Accu
racy = 0.8695878982543945
Iter #2904064:  Learning rate = 0.001531:   Batch Loss = 0.697637, Accurac
y = 0.869140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7208327054977417, Accu
racy = 0.8607198596000671
Iter #2908160:  Learning rate = 0.001531:   Batch Loss = 0.700716, Accurac
y = 0.8671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7302581071853638, Accu
racy = 0.8680229783058167
Iter #2912256:  Learning rate = 0.001531:   Batch Loss = 0.714719, Accurac
y = 0.857421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7125714421272278, Accu
racy = 0.8702834248542786
Iter #2916352:  Learning rate = 0.001531:   Batch Loss = 0.667261, Accurac
y = 0.8828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7157679200172424, Accu
racy = 0.8551556468009949
Iter #2920448:  Learning rate = 0.001531:   Batch Loss = 0.643620, Accurac
y = 0.890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6781882047653198, Accu
racy = 0.8887150287628174
Iter #2924544:  Learning rate = 0.001531:   Batch Loss = 0.677344, Accurac
y = 0.884765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.701819121837616, Accur
acy = 0.8754999041557312
Iter #2928640:  Learning rate = 0.001531:   Batch Loss = 0.651428, Accurac
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7541266679763794, Accu
racy = 0.8615893125534058
Iter #2932736:  Learning rate = 0.001531:   Batch Loss = 0.644461, Accurac
y = 0.88671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7939023971557617, Accu
```

```
racy = 0.8412449955940247
Iter #2936832:  Learning rate = 0.001531:   Batch Loss = 0.608174, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.723032534122467, Accur
acy = 0.8577638864517212
Iter #2940928:  Learning rate = 0.001531:   Batch Loss = 0.654331, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6914047598838806, Accu
racy = 0.8836724162101746
Iter #2945024:  Learning rate = 0.001531:   Batch Loss = 0.703979, Accurac
y = 0.865234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6923818588256836, Accu
racy = 0.8661102652549744
Iter #2949120:  Learning rate = 0.001531:   Batch Loss = 0.584213, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6794596910476685, Accu
racy = 0.8845418095588684
Iter #2953216:  Learning rate = 0.001531:   Batch Loss = 0.660517, Accurac
y = 0.86328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6732302904129028, Accu
racy = 0.8901060819625854
Iter #2957312:  Learning rate = 0.001531:   Batch Loss = 0.587530, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7052233219146729, Accu
racy = 0.8753260374069214
Iter #2961408:  Learning rate = 0.001531:   Batch Loss = 0.631196, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6886124014854431, Accu
racy = 0.8848896026611328
Iter #2965504:  Learning rate = 0.001531:   Batch Loss = 0.607437, Accurac
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6773124933242798, Accu
racy = 0.8855851292610168
Iter #2969600:  Learning rate = 0.001531:   Batch Loss = 0.590208, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6781336665153503, Accu
racy = 0.8841940760612488
Iter #2973696:  Learning rate = 0.001531:   Batch Loss = 0.610695, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6726720333099365, Accu
racy = 0.8721961379051208
Iter #2977792:  Learning rate = 0.001531:   Batch Loss = 0.590008, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.665503740310669, Accur
acy = 0.8760215640068054
Iter #2981888:  Learning rate = 0.001531:   Batch Loss = 0.699020, Accurac
y = 0.8671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6813334226608276, Accu
racy = 0.8873239159584045
Iter #2985984:  Learning rate = 0.001531:   Batch Loss = 0.643322, Accurac
y = 0.869140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.698880672454834, Accur
acy = 0.857242226600647
Iter #2990080:  Learning rate = 0.001531:   Batch Loss = 0.620385, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.712397038936615, Accur
acy = 0.8735871911048889
Iter #2994176:  Learning rate = 0.001531:   Batch Loss = 0.677789, Accurac
y = 0.873046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6673047542572021, Accu
racy = 0.8786298036575317
```

```
Iter #2998272:  Learning rate = 0.001531:   Batch Loss = 0.695654, Accurac
y = 0.8671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7736760973930359, Accu
racy = 0.8283776640892029
Iter #3002368:  Learning rate = 0.001469:   Batch Loss = 0.731571, Accurac
y = 0.87109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6940262913703918, Accu
racy = 0.8768909573554993
Iter #3006464:  Learning rate = 0.001469:   Batch Loss = 0.610301, Accurac
y = 0.908203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.670367956161499, Accur
acy = 0.8786298036575317
Iter #3010560:  Learning rate = 0.001469:   Batch Loss = 0.557367, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6446675062179565, Accu
racy = 0.8958442211151123
Iter #3014656:  Learning rate = 0.001469:   Batch Loss = 0.647431, Accurac
y = 0.8828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7278150916099548, Accu
racy = 0.8542861938476562
Iter #3018752:  Learning rate = 0.001469:   Batch Loss = 0.615454, Accurac
y = 0.8984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6693516373634338, Accu
racy = 0.8838462829589844
Iter #3022848:  Learning rate = 0.001469:   Batch Loss = 0.652040, Accurac
y = 0.87890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6797966957092285, Accu
racy = 0.8840201497077942
Iter #3026944:  Learning rate = 0.001469:   Batch Loss = 0.617167, Accurac
y = 0.884765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.716065526008606, Accur
acy = 0.8603721261024475
Iter #3031040:  Learning rate = 0.001469:   Batch Loss = 0.686453, Accurac
y = 0.87890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6666976809501648, Accu
racy = 0.8914971351623535
Iter #3035136:  Learning rate = 0.001469:   Batch Loss = 0.581566, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6605812311172485, Accu
racy = 0.8930620551109314
Iter #3039232:  Learning rate = 0.001469:   Batch Loss = 0.688416, Accurac
y = 0.888671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6910697221755981, Accu
racy = 0.8765432238578796
Iter #3043328:  Learning rate = 0.001469:   Batch Loss = 0.616279, Accurac
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.668016254901886, Accur
acy = 0.8786298036575317
Iter #3047424:  Learning rate = 0.001469:   Batch Loss = 0.620258, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.658342719078064, Accur
acy = 0.8948009014129639
Iter #3051520:  Learning rate = 0.001469:   Batch Loss = 0.621266, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6485791206359863, Accu
racy = 0.8873239159584045
Iter #3055616:  Learning rate = 0.001469:   Batch Loss = 0.590867, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6504862308502197, Accu
racy = 0.8862806558609009
Iter #3059712:  Learning rate = 0.001469:   Batch Loss = 0.599093, Accurac
```

```
y = 0.89453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6405001878738403, Accu
racy = 0.8970614075660706
Iter #3063808:  Learning rate = 0.001469:   Batch Loss = 0.575355, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6527310013771057, Accu
racy = 0.8979308009147644
Iter #3067904:  Learning rate = 0.001469:   Batch Loss = 0.758795, Accurac
y = 0.853515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7063064575195312, Accu
racy = 0.870631217956543
Iter #3072000:  Learning rate = 0.001469:   Batch Loss = 0.651075, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6947566866874695, Accu
racy = 0.8688923716545105
Iter #3076096:  Learning rate = 0.001469:   Batch Loss = 0.806609, Accurac
y = 0.837890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7186765670776367, Accu
racy = 0.871326744556427
Iter #3080192:  Learning rate = 0.001469:   Batch Loss = 0.645827, Accurac
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7151709198951721, Accu
racy = 0.8741088509559631
Iter #3084288:  Learning rate = 0.001469:   Batch Loss = 0.681554, Accurac
y = 0.853515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7116309404373169, Accu
racy = 0.8504607677459717
Iter #3088384:  Learning rate = 0.001469:   Batch Loss = 0.639462, Accurac
y = 0.892578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6564059257507324, Accu
racy = 0.8880195021629333
Iter #3092480:  Learning rate = 0.001469:   Batch Loss = 0.598931, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6588230133056641, Accu
racy = 0.8874978423118591
Iter #3096576:  Learning rate = 0.001469:   Batch Loss = 0.568043, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6536502242088318, Accu
racy = 0.8880195021629333
Iter #3100672:  Learning rate = 0.001411:   Batch Loss = 0.609061, Accurac
y = 0.908203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6930086612701416, Accu
racy = 0.8812380433082581
Iter #3104768:  Learning rate = 0.001411:   Batch Loss = 0.626950, Accurac
y = 0.8828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6681365966796875, Accu
racy = 0.8688923716545105
Iter #3108864:  Learning rate = 0.001411:   Batch Loss = 0.598290, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6515148878097534, Accu
racy = 0.8921926617622375
Iter #3112960:  Learning rate = 0.001411:   Batch Loss = 0.548740, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.667192816734314, Accur
acy = 0.8899322152137756
Iter #3117056:  Learning rate = 0.001411:   Batch Loss = 0.646171, Accurac
y = 0.892578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6605244874954224, Accu
racy = 0.8788036704063416
Iter #3121152:  Learning rate = 0.001411:   Batch Loss = 0.592917, Accurac
y = 0.923828125
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.6700097322463989, Accu
racy = 0.8902799487113953
Iter #3125248:  Learning rate = 0.001411:   Batch Loss = 0.613009, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6752814054489136, Accu
racy = 0.8873239159584045
Iter #3129344:  Learning rate = 0.001411:   Batch Loss = 0.588996, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6679282188415527, Accu
racy = 0.8748043775558472
Iter #3133440:  Learning rate = 0.001411:   Batch Loss = 0.623284, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6625180840492249, Accu
racy = 0.8904538154602051
Iter #3137536:  Learning rate = 0.001411:   Batch Loss = 0.621729, Accurac
y = 0.892578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6702600717544556, Accu
racy = 0.8871500492095947
Iter #3141632:  Learning rate = 0.001411:   Batch Loss = 0.610044, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6586799621582031, Accu
racy = 0.89393150806427
Iter #3145728:  Learning rate = 0.001411:   Batch Loss = 0.596891, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6489585638046265, Accu
racy = 0.8822813630104065
Iter #3149824:  Learning rate = 0.001411:   Batch Loss = 0.617947, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6452418565750122, Accu
racy = 0.8993218541145325
Iter #3153920:  Learning rate = 0.001411:   Batch Loss = 0.706764, Accurac
y = 0.865234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6648589968681335, Accu
racy = 0.879151463508606
Iter #3158016:  Learning rate = 0.001411:   Batch Loss = 0.661114, Accurac
y = 0.873046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6708112955093384, Accu
racy = 0.8873239159584045
Iter #3162112:  Learning rate = 0.001411:   Batch Loss = 0.627954, Accurac
y = 0.8984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6397759318351746, Accu
racy = 0.8928881883621216
Iter #3166208:  Learning rate = 0.001411:   Batch Loss = 0.599669, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6571735143661499, Accu
racy = 0.8779342770576477
Iter #3170304:  Learning rate = 0.001411:   Batch Loss = 0.600374, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6403746008872986, Accu
racy = 0.8953225612640381
Iter #3174400:  Learning rate = 0.001411:   Batch Loss = 0.596623, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6496708393096924, Accu
racy = 0.8765432238578796
Iter #3178496:  Learning rate = 0.001411:   Batch Loss = 0.612077, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.666691243648529, Accur
acy = 0.8819335699081421
Iter #3182592:  Learning rate = 0.001411:   Batch Loss = 0.564816, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7382736802101135, Accu
```

```
racy = 0.8528951406478882
Iter #3186688:  Learning rate = 0.001411:   Batch Loss = 0.601975, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6913014650344849, Accu
racy = 0.8662841320037842
Iter #3190784:  Learning rate = 0.001411:   Batch Loss = 0.650918, Accurac
y = 0.873046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6596813201904297, Accu
racy = 0.8838462829589844
Iter #3194880:  Learning rate = 0.001411:   Batch Loss = 0.574858, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6404714584350586, Accu
racy = 0.8850634694099426
Iter #3198976:  Learning rate = 0.001411:   Batch Loss = 0.566849, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6365432739257812, Accu
racy = 0.896887481212616
Iter #3203072:  Learning rate = 0.001354:   Batch Loss = 0.650956, Accurac
y = 0.880859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6494165062904358, Accu
racy = 0.8914971351623535
Iter #3207168:  Learning rate = 0.001354:   Batch Loss = 0.656143, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6887513995170593, Accu
racy = 0.8767170906066895
Iter #3211264:  Learning rate = 0.001354:   Batch Loss = 0.596079, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6362704038619995, Accu
racy = 0.8954964280128479
Iter #3215360:  Learning rate = 0.001354:   Batch Loss = 0.592383, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6471551060676575, Accu
racy = 0.8775864839553833
Iter #3219456:  Learning rate = 0.001354:   Batch Loss = 0.601404, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6413225531578064, Accu
racy = 0.8831507563591003
Iter #3223552:  Learning rate = 0.001354:   Batch Loss = 0.535623, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6466091871261597, Accu
racy = 0.8920187950134277
Iter #3227648:  Learning rate = 0.001354:   Batch Loss = 0.570983, Accurac
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6396497488021851, Accu
racy = 0.8868023157119751
Iter #3231744:  Learning rate = 0.001354:   Batch Loss = 0.562597, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6437282562255859, Accu
racy = 0.8951486945152283
Iter #3235840:  Learning rate = 0.001354:   Batch Loss = 0.551507, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6552374362945557, Accu
racy = 0.8786298036575317
Iter #3239936:  Learning rate = 0.001354:   Batch Loss = 0.589840, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6518359184265137, Accu
racy = 0.8796731233596802
Iter #3244032:  Learning rate = 0.001354:   Batch Loss = 0.548762, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6387667655944824, Accu
racy = 0.894627034664154
```

```
Iter #3248128:  Learning rate = 0.001354:   Batch Loss = 0.549955, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6416323184967041, Accu
racy = 0.8895844221115112
Iter #3252224:  Learning rate = 0.001354:   Batch Loss = 0.600841, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6727404594421387, Accu
racy = 0.8753260374069214
Iter #3256320:  Learning rate = 0.001354:   Batch Loss = 0.536734, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6557528972625732, Accu
racy = 0.8930620551109314
Iter #3260416:  Learning rate = 0.001354:   Batch Loss = 0.574456, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6485773324966431, Accu
racy = 0.893235981464386
Iter #3264512:  Learning rate = 0.001354:   Batch Loss = 0.603581, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6828123927116394, Accu
racy = 0.8814119100570679
Iter #3268608:  Learning rate = 0.001354:   Batch Loss = 0.652878, Accurac
y = 0.880859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6635400056838989, Accu
racy = 0.8812380433082581
Iter #3272704:  Learning rate = 0.001354:   Batch Loss = 0.694872, Accurac
y = 0.84765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6503121852874756, Accu
racy = 0.8796731233596802
Iter #3276800:  Learning rate = 0.001354:   Batch Loss = 0.583710, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6425055861473083, Accu
racy = 0.8967136144638062
Iter #3280896:  Learning rate = 0.001354:   Batch Loss = 0.556750, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6421613097190857, Accu
racy = 0.8965397477149963
Iter #3284992:  Learning rate = 0.001354:   Batch Loss = 0.593310, Accurac
y = 0.908203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6605731844902039, Accu
racy = 0.8836724162101746
Iter #3289088:  Learning rate = 0.001354:   Batch Loss = 0.582187, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6731743812561035, Accu
racy = 0.8732394576072693
Iter #3293184:  Learning rate = 0.001354:   Batch Loss = 0.576105, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6466308832168579, Accu
racy = 0.8918448686599731
Iter #3297280:  Learning rate = 0.001354:   Batch Loss = 0.560798, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6683669686317444, Accu
racy = 0.8751521706581116
Iter #3301376:  Learning rate = 0.001300:   Batch Loss = 0.584661, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6446294784545898, Accu
racy = 0.8821074366569519
Iter #3305472:  Learning rate = 0.001300:   Batch Loss = 0.631879, Accurac
y = 0.89453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.628947913646698, Accur
acy = 0.888367235660553
Iter #3309568:  Learning rate = 0.001300:   Batch Loss = 0.562689, Accurac
```

```
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6303668022155762, Accu
racy = 0.8998435139656067
Iter #3313664:  Learning rate = 0.001300:   Batch Loss = 0.589085, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6208632588386536, Accu
racy = 0.9008867740631104
Iter #3317760:  Learning rate = 0.001300:   Batch Loss = 0.562610, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6296918392181396, Accu
racy = 0.8881933689117432
Iter #3321856:  Learning rate = 0.001300:   Batch Loss = 0.614780, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6543347835540771, Accu
racy = 0.8819335699081421
Iter #3325952:  Learning rate = 0.001300:   Batch Loss = 0.593154, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.647034764289856, Accur
acy = 0.892540454864502
Iter #3330048:  Learning rate = 0.001300:   Batch Loss = 0.588262, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6479647159576416, Accu
racy = 0.8833246231079102
Iter #3334144:  Learning rate = 0.001300:   Batch Loss = 0.560451, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6377781629562378, Accu
racy = 0.9007129073143005
Iter #3338240:  Learning rate = 0.001300:   Batch Loss = 0.562615, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6639331579208374, Accu
racy = 0.874282717704773
Iter #3342336:  Learning rate = 0.001300:   Batch Loss = 0.607120, Accurac
y = 0.890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6324560046195984, Accu
racy = 0.8974091410636902
Iter #3346432:  Learning rate = 0.001300:   Batch Loss = 0.577593, Accurac
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6361640691757202, Accu
racy = 0.8961919546127319
Iter #3350528:  Learning rate = 0.001300:   Batch Loss = 0.601940, Accurac
y = 0.908203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.621391773223877, Accur
acy = 0.8935837149620056
Iter #3354624:  Learning rate = 0.001300:   Batch Loss = 0.544772, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6388837099075317, Accu
racy = 0.8881933689117432
Iter #3358720:  Learning rate = 0.001300:   Batch Loss = 0.593740, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6316921710968018, Accu
racy = 0.8817597031593323
Iter #3362816:  Learning rate = 0.001300:   Batch Loss = 0.598449, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6209471821784973, Accu
racy = 0.8857589960098267
Iter #3366912:  Learning rate = 0.001300:   Batch Loss = 0.566896, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6329228281974792, Accu
racy = 0.8948009014129639
Iter #3371008:  Learning rate = 0.001300:   Batch Loss = 0.585629, Accurac
y = 0.916015625
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.6471407413482666, Accu
racy = 0.8834985494613647
Iter #3375104:  Learning rate = 0.001300:   Batch Loss = 0.542030, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6329561471939087, Accu
racy = 0.8850634694099426
Iter #3379200:  Learning rate = 0.001300:   Batch Loss = 0.551697, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6218336820602417, Accu
racy = 0.9026256203651428
Iter #3383296:  Learning rate = 0.001300:   Batch Loss = 0.532236, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6305158138275146, Accu
racy = 0.8894105553627014
Iter #3387392:  Learning rate = 0.001300:   Batch Loss = 0.565974, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6212453842163086, Accu
racy = 0.892540454864502
Iter #3391488:  Learning rate = 0.001300:   Batch Loss = 0.574425, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6583229303359985, Accu
racy = 0.874282717704773
Iter #3395584:  Learning rate = 0.001300:   Batch Loss = 0.563508, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6359311938285828, Accu
racy = 0.8887150287628174
Iter #3399680:  Learning rate = 0.001300:   Batch Loss = 0.613886, Accurac
y = 0.892578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7038145065307617, Accu
racy = 0.8678490519523621
Iter #3403776:  Learning rate = 0.001248:   Batch Loss = 0.587374, Accurac
y = 0.890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6676182746887207, Accu
racy = 0.8869761824607849
Iter #3407872:  Learning rate = 0.001248:   Batch Loss = 0.625176, Accurac
y = 0.884765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.659644603729248, Accur
acy = 0.8725439310073853
Iter #3411968:  Learning rate = 0.001248:   Batch Loss = 0.609398, Accurac
y = 0.888671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6357436776161194, Accu
racy = 0.8941053748130798
Iter #3416064:  Learning rate = 0.001248:   Batch Loss = 0.779501, Accurac
y = 0.83984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.750482439994812, Accur
acy = 0.8546339869499207
Iter #3420160:  Learning rate = 0.001248:   Batch Loss = 0.608332, Accurac
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6805396676063538, Accu
racy = 0.8770648837089539
Iter #3424256:  Learning rate = 0.001248:   Batch Loss = 0.611991, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6931196451187134, Accu
racy = 0.8748043775558472
Iter #3428352:  Learning rate = 0.001248:   Batch Loss = 0.573308, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6722413897514343, Accu
racy = 0.8772387504577637
Iter #3432448:  Learning rate = 0.001248:   Batch Loss = 0.619504, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6610772609710693, Accu
```

```
racy = 0.8829768896102905
Iter #3436544:  Learning rate = 0.001248:   Batch Loss = 0.565938, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6413213014602661, Accu
racy = 0.8937575817108154
Iter #3440640:  Learning rate = 0.001248:   Batch Loss = 0.536671, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6173089742660522, Accu
racy = 0.9010607004165649
Iter #3444736:  Learning rate = 0.001248:   Batch Loss = 0.558511, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6000082492828369, Accu
racy = 0.9083637595176697
Iter #3448832:  Learning rate = 0.001248:   Batch Loss = 0.533848, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6130942106246948, Accu
racy = 0.9010607004165649
Iter #3452928:  Learning rate = 0.001248:   Batch Loss = 0.610295, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6177054643630981, Accu
racy = 0.9040166735649109
Iter #3457024:  Learning rate = 0.001248:   Batch Loss = 0.523855, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6376559734344482, Accu
racy = 0.8970614075660706
Iter #3461120:  Learning rate = 0.001248:   Batch Loss = 0.624972, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6196956634521484, Accu
racy = 0.8930620551109314
Iter #3465216:  Learning rate = 0.001248:   Batch Loss = 0.560938, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6400959491729736, Accu
racy = 0.8951486945152283
Iter #3469312:  Learning rate = 0.001248:   Batch Loss = 0.556367, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6136130094528198, Accu
racy = 0.9001912474632263
Iter #3473408:  Learning rate = 0.001248:   Batch Loss = 0.562275, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6096651554107666, Accu
racy = 0.9057555198669434
Iter #3477504:  Learning rate = 0.001248:   Batch Loss = 0.581637, Accurac
y = 0.908203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6071568727493286, Accu
racy = 0.8967136144638062
Iter #3481600:  Learning rate = 0.001248:   Batch Loss = 0.562791, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6201263666152954, Accu
racy = 0.9022778868675232
Iter #3485696:  Learning rate = 0.001248:   Batch Loss = 0.538750, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6415537595748901, Accu
racy = 0.901756227016449
Iter #3489792:  Learning rate = 0.001248:   Batch Loss = 0.518808, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5951422452926636, Accu
racy = 0.9113197922706604
Iter #3493888:  Learning rate = 0.001248:   Batch Loss = 0.551998, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5969562530517578, Accu
racy = 0.9118413925170898
```

```
Iter #3497984:  Learning rate = 0.001248:   Batch Loss = 0.541729, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6088017225265503, Accu
racy = 0.9090592861175537
Iter #3502080:  Learning rate = 0.001198:   Batch Loss = 0.524883, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5958379507064819, Accu
racy = 0.9146235585212708
Iter #3506176:  Learning rate = 0.001198:   Batch Loss = 0.562373, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6040958166122437, Accu
racy = 0.9055816531181335
Iter #3510272:  Learning rate = 0.001198:   Batch Loss = 0.522835, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5905100703239441, Accu
racy = 0.910971999168396
Iter #3514368:  Learning rate = 0.001198:   Batch Loss = 0.494668, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6125175952911377, Accu
racy = 0.9099286794662476
Iter #3518464:  Learning rate = 0.001198:   Batch Loss = 0.559716, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5959605574607849, Accu
racy = 0.9062771797180176
Iter #3522560:  Learning rate = 0.001198:   Batch Loss = 0.567846, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6067091226577759, Accu
racy = 0.9074943661689758
Iter #3526656:  Learning rate = 0.001198:   Batch Loss = 0.568860, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5890787243843079, Accu
racy = 0.915145218372345
Iter #3530752:  Learning rate = 0.001198:   Batch Loss = 0.516215, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5815877318382263, Accu
racy = 0.9114936590194702
Iter #3534848:  Learning rate = 0.001198:   Batch Loss = 0.567066, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6006016731262207, Accu
racy = 0.9092331528663635
Iter #3538944:  Learning rate = 0.001198:   Batch Loss = 0.526356, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5860342979431152, Accu
racy = 0.9099286794662476
Iter #3543040:  Learning rate = 0.001198:   Batch Loss = 0.511477, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.63090980052948, Accura
cy = 0.8963658213615417
Iter #3547136:  Learning rate = 0.001198:   Batch Loss = 0.519083, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5975768566131592, Accu
racy = 0.9069727063179016
Iter #3551232:  Learning rate = 0.001198:   Batch Loss = 0.494240, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5905280113220215, Accu
racy = 0.9125369787216187
Iter #3555328:  Learning rate = 0.001198:   Batch Loss = 0.497472, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6003861427307129, Accu
racy = 0.91166752576828
Iter #3559424:  Learning rate = 0.001198:   Batch Loss = 0.540215, Accurac
```

```
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5938690304756165, Accu
racy = 0.9114936590194702
Iter #3563520:  Learning rate = 0.001198:   Batch Loss = 0.528473, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6222648620605469, Accu
racy = 0.9064510464668274
Iter #3567616:  Learning rate = 0.001198:   Batch Loss = 0.520224, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5958450436592102, Accu
racy = 0.9127108454704285
Iter #3571712:  Learning rate = 0.001198:   Batch Loss = 0.548229, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5895788669586182, Accu
racy = 0.9101026058197021
Iter #3575808:  Learning rate = 0.001198:   Batch Loss = 0.511692, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5851582288742065, Accu
racy = 0.9167101383209229
Iter #3579904:  Learning rate = 0.001198:   Batch Loss = 0.518421, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5867494344711304, Accu
racy = 0.9181011915206909
Iter #3584000:  Learning rate = 0.001198:   Batch Loss = 0.504387, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.608558177947998, Accur
acy = 0.9099286794662476
Iter #3588096:  Learning rate = 0.001198:   Batch Loss = 0.570908, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6292814016342163, Accu
racy = 0.8988001942634583
Iter #3592192:  Learning rate = 0.001198:   Batch Loss = 0.601540, Accurac
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6189349293708801, Accu
racy = 0.8941053748130798
Iter #3596288:  Learning rate = 0.001198:   Batch Loss = 0.534113, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6127307415008545, Accu
racy = 0.9005390405654907
Iter #3600384:  Learning rate = 0.001150:   Batch Loss = 0.538709, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6147638559341431, Accu
racy = 0.9066249132156372
Iter #3604480:  Learning rate = 0.001150:   Batch Loss = 0.553849, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6287510395050049, Accu
racy = 0.8894105553627014
Iter #3608576:  Learning rate = 0.001150:   Batch Loss = 0.528147, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5738019943237305, Accu
racy = 0.9134063720703125
Iter #3612672:  Learning rate = 0.001150:   Batch Loss = 0.561525, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5991308689117432, Accu
racy = 0.9083637595176697
Iter #3616768:  Learning rate = 0.001150:   Batch Loss = 0.558849, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6051582098007202, Accu
racy = 0.9033211469650269
Iter #3620864:  Learning rate = 0.001150:   Batch Loss = 0.551311, Accurac
y = 0.9375
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.6122322082519531, Accu
racy = 0.9104503393173218
Iter #3624960:  Learning rate = 0.001150:   Batch Loss = 0.509679, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5988365411758423, Accu
racy = 0.9088854193687439
Iter #3629056:  Learning rate = 0.001150:   Batch Loss = 0.520136, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.588845431804657, Accur
acy = 0.9099286794662476
Iter #3633152:  Learning rate = 0.001150:   Batch Loss = 0.578997, Accurac
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6392368078231812, Accu
racy = 0.8948009014129639
Iter #3637248:  Learning rate = 0.001150:   Batch Loss = 0.612010, Accurac
y = 0.888671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6312434673309326, Accu
racy = 0.8824552297592163
Iter #3641344:  Learning rate = 0.001150:   Batch Loss = 0.574883, Accurac
y = 0.89453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5918912291526794, Accu
racy = 0.9036689400672913
Iter #3645440:  Learning rate = 0.001150:   Batch Loss = 0.540884, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5870991945266724, Accu
racy = 0.9095809459686279
Iter #3649536:  Learning rate = 0.001150:   Batch Loss = 0.529461, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5834629535675049, Accu
racy = 0.9083637595176697
Iter #3653632:  Learning rate = 0.001150:   Batch Loss = 0.528316, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5668783187866211, Accu
racy = 0.9226221442222595
Iter #3657728:  Learning rate = 0.001150:   Batch Loss = 0.484939, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5743100643157959, Accu
racy = 0.9121891856193542
Iter #3661824:  Learning rate = 0.001150:   Batch Loss = 0.515388, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5789372324943542, Accu
racy = 0.9094070792198181
Iter #3665920:  Learning rate = 0.001150:   Batch Loss = 0.528574, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5716521143913269, Accu
racy = 0.9198400378227234
Iter #3670016:  Learning rate = 0.001150:   Batch Loss = 0.581668, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.578394889831543, Accur
acy = 0.9113197922706604
Iter #3674112:  Learning rate = 0.001150:   Batch Loss = 0.569655, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6115889549255371, Accu
racy = 0.9066249132156372
Iter #3678208:  Learning rate = 0.001150:   Batch Loss = 0.540160, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5866207480430603, Accu
racy = 0.9121891856193542
Iter #3682304:  Learning rate = 0.001150:   Batch Loss = 0.524880, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5885761976242065, Accu
```

```
racy = 0.9064510464668274
Iter #3686400:  Learning rate = 0.001150:   Batch Loss = 0.481743, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5825322270393372, Accu
racy = 0.9085376262664795
Iter #3690496:  Learning rate = 0.001150:   Batch Loss = 0.518684, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5942084789276123, Accu
racy = 0.91166752576828
Iter #3694592:  Learning rate = 0.001150:   Batch Loss = 0.548425, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5760058164596558, Accu
racy = 0.9113197922706604
Iter #3698688:  Learning rate = 0.001150:   Batch Loss = 0.516389, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5870729684829712, Accu
racy = 0.9134063720703125
Iter #3702784:  Learning rate = 0.001104:   Batch Loss = 0.514676, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5658901333808899, Accu
racy = 0.9168840050697327
Iter #3706880:  Learning rate = 0.001104:   Batch Loss = 0.488470, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5806888341903687, Accu
racy = 0.90801602602005
Iter #3710976:  Learning rate = 0.001104:   Batch Loss = 0.558738, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6116983294487, Accurac
y = 0.8942792415618896
Iter #3715072:  Learning rate = 0.001104:   Batch Loss = 0.532282, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5988644361495972, Accu
racy = 0.9062771797180176
Iter #3719168:  Learning rate = 0.001104:   Batch Loss = 0.539142, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6088854074478149, Accu
racy = 0.9083637595176697
Iter #3723264:  Learning rate = 0.001104:   Batch Loss = 0.522955, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5804195404052734, Accu
racy = 0.9184489846229553
Iter #3727360:  Learning rate = 0.001104:   Batch Loss = 0.530611, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5749397277832031, Accu
racy = 0.9114936590194702
Iter #3731456:  Learning rate = 0.001104:   Batch Loss = 0.552105, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6002840995788574, Accu
racy = 0.9062771797180176
Iter #3735552:  Learning rate = 0.001104:   Batch Loss = 0.499460, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5976368188858032, Accu
racy = 0.9156668186187744
Iter #3739648:  Learning rate = 0.001104:   Batch Loss = 0.513544, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5803556442260742, Accu
racy = 0.9142757654190063
Iter #3743744:  Learning rate = 0.001104:   Batch Loss = 0.479956, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5671409368515015, Accu
racy = 0.919492244720459
```

```
Iter #3747840:  Learning rate = 0.001104:   Batch Loss = 0.526270, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5896515846252441, Accu
racy = 0.9154929518699646
Iter #3751936:  Learning rate = 0.001104:   Batch Loss = 0.468789, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5843260288238525, Accu
racy = 0.9149712920188904
Iter #3756032:  Learning rate = 0.001104:   Batch Loss = 0.576196, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5975240468978882, Accu
racy = 0.8986263275146484
Iter #3760128:  Learning rate = 0.001104:   Batch Loss = 0.523905, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6017062664031982, Accu
racy = 0.894627034664154
Iter #3764224:  Learning rate = 0.001104:   Batch Loss = 0.515790, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6019786596298218, Accu
racy = 0.8944531679153442
Iter #3768320:  Learning rate = 0.001104:   Batch Loss = 0.536992, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5811336040496826, Accu
racy = 0.910276472568512
Iter #3772416:  Learning rate = 0.001104:   Batch Loss = 0.537337, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5919690132141113, Accu
racy = 0.9088854193687439
Iter #3776512:  Learning rate = 0.001104:   Batch Loss = 0.517498, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5983930230140686, Accu
racy = 0.9146235585212708
Iter #3780608:  Learning rate = 0.001104:   Batch Loss = 0.594472, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6237317323684692, Accu
racy = 0.8862806558609009
Iter #3784704:  Learning rate = 0.001104:   Batch Loss = 0.465216, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5680892467498779, Accu
racy = 0.9113197922706604
Iter #3788800:  Learning rate = 0.001104:   Batch Loss = 0.455245, Accurac
y = 0.962890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5651236772537231, Accu
racy = 0.9186228513717651
Iter #3792896:  Learning rate = 0.001104:   Batch Loss = 0.500486, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5811206102371216, Accu
racy = 0.9137541055679321
Iter #3796992:  Learning rate = 0.001104:   Batch Loss = 0.480343, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5588409900665283, Accu
racy = 0.9191445112228394
Iter #3801088:  Learning rate = 0.001060:   Batch Loss = 0.508270, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5717681646347046, Accu
racy = 0.9186228513717651
Iter #3805184:  Learning rate = 0.001060:   Batch Loss = 0.483906, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5916274785995483, Accu
racy = 0.9121891856193542
Iter #3809280:  Learning rate = 0.001060:   Batch Loss = 0.483687, Accurac
```

```
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5855780839920044, Accu
racy = 0.9141018986701965
Iter #3813376:  Learning rate = 0.001060:   Batch Loss = 0.477056, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5660417079925537, Accu
racy = 0.9215788841247559
Iter #3817472:  Learning rate = 0.001060:   Batch Loss = 0.519935, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5952285528182983, Accu
racy = 0.9127108454704285
Iter #3821568:  Learning rate = 0.001060:   Batch Loss = 0.542201, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5817620754241943, Accu
racy = 0.910971999168396
Iter #3825664:  Learning rate = 0.001060:   Batch Loss = 0.557058, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6139352321624756, Accu
racy = 0.9027994871139526
Iter #3829760:  Learning rate = 0.001060:   Batch Loss = 0.506086, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5777825117111206, Accu
racy = 0.9147974252700806
Iter #3833856:  Learning rate = 0.001060:   Batch Loss = 0.473221, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5920391082763672, Accu
racy = 0.9067988395690918
Iter #3837952:  Learning rate = 0.001060:   Batch Loss = 0.518720, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5766373872756958, Accu
racy = 0.9106242656707764
Iter #3842048:  Learning rate = 0.001060:   Batch Loss = 0.490712, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5755591988563538, Accu
racy = 0.9160146117210388
Iter #3846144:  Learning rate = 0.001060:   Batch Loss = 0.525251, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5807737708091736, Accu
racy = 0.9052338600158691
Iter #3850240:  Learning rate = 0.001060:   Batch Loss = 0.512126, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5840578675270081, Accu
racy = 0.9196661710739136
Iter #3854336:  Learning rate = 0.001060:   Batch Loss = 0.472266, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5692307949066162, Accu
racy = 0.9172317981719971
Iter #3858432:  Learning rate = 0.001060:   Batch Loss = 0.494631, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5664543509483337, Accu
racy = 0.9128847122192383
Iter #3862528:  Learning rate = 0.001060:   Batch Loss = 0.512182, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5696609020233154, Accu
racy = 0.9121891856193542
Iter #3866624:  Learning rate = 0.001060:   Batch Loss = 0.489612, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5719882845878601, Accu
racy = 0.9146235585212708
Iter #3870720:  Learning rate = 0.001060:   Batch Loss = 0.526575, Accurac
y = 0.9375
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.5793415904045105, Accu
racy = 0.9114936590194702
Iter #3874816:  Learning rate = 0.001060:   Batch Loss = 0.490698, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5728042125701904, Accu
racy = 0.9149712920188904
Iter #3878912:  Learning rate = 0.001060:   Batch Loss = 0.523027, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5865359902381897, Accu
racy = 0.915840744972229
Iter #3883008:  Learning rate = 0.001060:   Batch Loss = 0.500639, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5680481791496277, Accu
racy = 0.9127108454704285
Iter #3887104:  Learning rate = 0.001060:   Batch Loss = 0.483208, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5643750429153442, Accu
racy = 0.9167101383209229
Iter #3891200:  Learning rate = 0.001060:   Batch Loss = 0.490765, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5704269409179688, Accu
racy = 0.9137541055679321
Iter #3895296:  Learning rate = 0.001060:   Batch Loss = 0.464651, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5708049535751343, Accu
racy = 0.910971999168396
Iter #3899392:  Learning rate = 0.001060:   Batch Loss = 0.473993, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5923012495040894, Accu
racy = 0.9048861265182495
Iter #3903488:  Learning rate = 0.001018:   Batch Loss = 0.509412, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5895946025848389, Accu
racy = 0.9066249132156372
Iter #3907584:  Learning rate = 0.001018:   Batch Loss = 0.488422, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5835488438606262, Accu
racy = 0.9132325053215027
Iter #3911680:  Learning rate = 0.001018:   Batch Loss = 0.554234, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6059345602989197, Accu
racy = 0.9067988395690918
Iter #3915776:  Learning rate = 0.001018:   Batch Loss = 0.522640, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6044797301292419, Accu
racy = 0.9022778868675232
Iter #3919872:  Learning rate = 0.001018:   Batch Loss = 0.502468, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5888908505439758, Accu
racy = 0.9036689400672913
Iter #3923968:  Learning rate = 0.001018:   Batch Loss = 0.540173, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6057010889053345, Accu
racy = 0.903147280216217
Iter #3928064:  Learning rate = 0.001018:   Batch Loss = 0.551580, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5667709112167358, Accu
racy = 0.9104503393173218
Iter #3932160:  Learning rate = 0.001018:   Batch Loss = 0.523546, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5925582051277161, Accu
```

```
racy = 0.9113197922706604
Iter #3936256:  Learning rate = 0.001018:   Batch Loss = 0.504860, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5824821591377258, Accu
racy = 0.9181011915206909
Iter #3940352:  Learning rate = 0.001018:   Batch Loss = 0.546157, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.581186830997467, Accur
acy = 0.9161884784698486
Iter #3944448:  Learning rate = 0.001018:   Batch Loss = 0.509426, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5603655576705933, Accu
racy = 0.9175795316696167
Iter #3948544:  Learning rate = 0.001018:   Batch Loss = 0.470632, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5676988363265991, Accu
racy = 0.916536271572113
Iter #3952640:  Learning rate = 0.001018:   Batch Loss = 0.505322, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5792475938796997, Accu
racy = 0.9167101383209229
Iter #3956736:  Learning rate = 0.001018:   Batch Loss = 0.501466, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5916540622711182, Accu
racy = 0.9144496321678162
Iter #3960832:  Learning rate = 0.001018:   Batch Loss = 0.478218, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5982575416564941, Accu
racy = 0.9111458659172058
Iter #3964928:  Learning rate = 0.001018:   Batch Loss = 0.470507, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5904977917671204, Accu
racy = 0.9074943661689758
Iter #3969024:  Learning rate = 0.001018:   Batch Loss = 0.467743, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6039277911186218, Accu
racy = 0.9043644666671753
Iter #3973120:  Learning rate = 0.001018:   Batch Loss = 0.514270, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5740876793861389, Accu
racy = 0.9127108454704285
Iter #3977216:  Learning rate = 0.001018:   Batch Loss = 0.488803, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5697535872459412, Accu
racy = 0.915840744972229
Iter #3981312:  Learning rate = 0.001018:   Batch Loss = 0.517587, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5509692430496216, Accu
racy = 0.9214049577713013
Iter #3985408:  Learning rate = 0.001018:   Batch Loss = 0.480215, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5704783201217651, Accu
racy = 0.9186228513717651
Iter #3989504:  Learning rate = 0.001018:   Batch Loss = 0.474661, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5625893473625183, Accu
racy = 0.9212310910224915
Iter #3993600:  Learning rate = 0.001018:   Batch Loss = 0.487104, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5552747249603271, Accu
racy = 0.915840744972229
```

```
Iter #3997696:  Learning rate = 0.001018:   Batch Loss = 0.517864, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5671321749687195, Accu
racy = 0.9168840050697327
Iter #4001792:  Learning rate = 0.000977:   Batch Loss = 0.496245, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5666284561157227, Accu
racy = 0.9241871237754822
Iter #4005888:  Learning rate = 0.000977:   Batch Loss = 0.469472, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5680243968963623, Accu
racy = 0.9179273247718811
Iter #4009984:  Learning rate = 0.000977:   Batch Loss = 0.514608, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5591322779655457, Accu
racy = 0.9219266176223755
Iter #4014080:  Learning rate = 0.000977:   Batch Loss = 0.513972, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5634260177612305, Accu
racy = 0.9149712920188904
Iter #4018176:  Learning rate = 0.000977:   Batch Loss = 0.506427, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5717043280601501, Accu
racy = 0.9252303838729858
Iter #4022272:  Learning rate = 0.000977:   Batch Loss = 0.491345, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.556774377822876, Accur
acy = 0.9141018986701965
Iter #4026368:  Learning rate = 0.000977:   Batch Loss = 0.511068, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5791484713554382, Accu
racy = 0.9074943661689758
Iter #4030464:  Learning rate = 0.000977:   Batch Loss = 0.456077, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6205611228942871, Accu
racy = 0.9034950733184814
Iter #4034560:  Learning rate = 0.000977:   Batch Loss = 0.515707, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.562319278717041, Accur
acy = 0.918796718120575
Iter #4038656:  Learning rate = 0.000977:   Batch Loss = 0.553960, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6033790111541748, Accu
racy = 0.8937575817108154
Iter #4042752:  Learning rate = 0.000977:   Batch Loss = 0.496978, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5645890235900879, Accu
racy = 0.915840744972229
Iter #4046848:  Learning rate = 0.000977:   Batch Loss = 0.492548, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5753213763237, Accurac
y = 0.9193183779716492
Iter #4050944:  Learning rate = 0.000977:   Batch Loss = 0.476997, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5650576949119568, Accu
racy = 0.9163623452186584
Iter #4055040:  Learning rate = 0.000977:   Batch Loss = 0.565195, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5744633674621582, Accu
racy = 0.9106242656707764
Iter #4059136:  Learning rate = 0.000977:   Batch Loss = 0.489146, Accurac
```

```
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5699996948242188, Accu
racy = 0.919492244720459
Iter #4063232:  Learning rate = 0.000977:   Batch Loss = 0.552475, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5783556699752808, Accu
racy = 0.9121891856193542
Iter #4067328:  Learning rate = 0.000977:   Batch Loss = 0.462858, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5792312622070312, Accu
racy = 0.9118413925170898
Iter #4071424:  Learning rate = 0.000977:   Batch Loss = 0.498048, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5627503395080566, Accu
racy = 0.9149712920188904
Iter #4075520:  Learning rate = 0.000977:   Batch Loss = 0.498695, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5593833923339844, Accu
racy = 0.9172317981719971
Iter #4079616:  Learning rate = 0.000977:   Batch Loss = 0.445152, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5573114156723022, Accu
racy = 0.9217527508735657
Iter #4083712:  Learning rate = 0.000977:   Batch Loss = 0.448480, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5485256910324097, Accu
racy = 0.920187771320343
Iter #4087808:  Learning rate = 0.000977:   Batch Loss = 0.482434, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5692552328109741, Accu
racy = 0.9221004843711853
Iter #4091904:  Learning rate = 0.000977:   Batch Loss = 0.493700, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5610412359237671, Accu
racy = 0.9245348572731018
Iter #4096000:  Learning rate = 0.000977:   Batch Loss = 0.461787, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5554791688919067, Accu
racy = 0.9170579314231873
Iter #4100096:  Learning rate = 0.000938:   Batch Loss = 0.496712, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5516250729560852, Accu
racy = 0.9219266176223755
Iter #4104192:  Learning rate = 0.000938:   Batch Loss = 0.443448, Accurac
y = 0.966796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5797591805458069, Accu
racy = 0.9141018986701965
Iter #4108288:  Learning rate = 0.000938:   Batch Loss = 0.494324, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5669769048690796, Accu
racy = 0.9142757654190063
Iter #4112384:  Learning rate = 0.000938:   Batch Loss = 0.525183, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5832394361495972, Accu
racy = 0.9057555198669434
Iter #4116480:  Learning rate = 0.000938:   Batch Loss = 0.511002, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5730434060096741, Accu
racy = 0.9179273247718811
Iter #4120576:  Learning rate = 0.000938:   Batch Loss = 0.457776, Accurac
y = 0.953125
```

```
PERFORMANCE ON TEST SET:            Batch Loss = 0.5678778886795044, Accu
racy = 0.9156668186187744
Iter #4124672:  Learning rate = 0.000938:   Batch Loss = 0.459090, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5642760992050171, Accu
racy = 0.9160146117210388
Iter #4128768:  Learning rate = 0.000938:   Batch Loss = 0.477837, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5582672357559204, Accu
racy = 0.9156668186187744
Iter #4132864:  Learning rate = 0.000938:   Batch Loss = 0.462331, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:            Batch Loss = 0.566373884677887, Accur
acy = 0.9177534580230713
Iter #4136960:  Learning rate = 0.000938:   Batch Loss = 0.519280, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:            Batch Loss = 0.554519534111023, Accur
acy = 0.9193183779716492
Iter #4141056:  Learning rate = 0.000938:   Batch Loss = 0.463639, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5544813871383667, Accu
racy = 0.9184489846229553
Iter #4145152:  Learning rate = 0.000938:   Batch Loss = 0.503228, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5601391792297363, Accu
racy = 0.9182750582695007
Iter #4149248:  Learning rate = 0.000938:   Batch Loss = 0.495492, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5759526491165161, Accu
racy = 0.9104503393173218
Iter #4153344:  Learning rate = 0.000938:   Batch Loss = 0.499364, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5844981074333191, Accu
racy = 0.9196661710739136
Iter #4157440:  Learning rate = 0.000938:   Batch Loss = 0.512110, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5866352319717407, Accu
racy = 0.9101026058197021
Iter #4161536:  Learning rate = 0.000938:   Batch Loss = 0.497277, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5842663049697876, Accu
racy = 0.9149712920188904
Iter #4165632:  Learning rate = 0.000938:   Batch Loss = 0.443613, Accurac
y = 0.962890625
PERFORMANCE ON TEST SET:            Batch Loss = 0.5652750134468079, Accu
racy = 0.9123630523681641
Iter #4169728:  Learning rate = 0.000938:   Batch Loss = 0.509349, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5558255910873413, Accu
racy = 0.9179273247718811
Iter #4173824:  Learning rate = 0.000938:   Batch Loss = 0.490383, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5595437288284302, Accu
racy = 0.9085376262664795
Iter #4177920:  Learning rate = 0.000938:   Batch Loss = 0.526735, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:            Batch Loss = 0.5878417491912842, Accu
racy = 0.9055816531181335
Iter #4182016:  Learning rate = 0.000938:   Batch Loss = 0.467744, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:            Batch Loss = 0.5583495497703552, Accu
```

```
racy = 0.9156668186187744
Iter #4186112:  Learning rate = 0.000938:   Batch Loss = 0.463898, Accurac
y = 0.962890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5593214631080627, Accu
racy = 0.9146235585212708
Iter #4190208:  Learning rate = 0.000938:   Batch Loss = 0.503346, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5539811849594116, Accu
racy = 0.9174056649208069
Iter #4194304:  Learning rate = 0.000938:   Batch Loss = 0.493113, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5574874877929688, Accu
racy = 0.9238393306732178
Iter #4198400:  Learning rate = 0.000938:   Batch Loss = 0.447862, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5538020133972168, Accu
racy = 0.9177534580230713
Iter #4202496:  Learning rate = 0.000900:   Batch Loss = 0.467529, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5383772253990173, Accu
racy = 0.9226221442222595
Iter #4206592:  Learning rate = 0.000900:   Batch Loss = 0.480320, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5558720231056213, Accu
racy = 0.9191445112228394
Iter #4210688:  Learning rate = 0.000900:   Batch Loss = 0.460343, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5680951476097107, Accu
racy = 0.9107981324195862
Iter #4214784:  Learning rate = 0.000900:   Batch Loss = 0.464386, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5647702217102051, Accu
racy = 0.9134063720703125
Iter #4218880:  Learning rate = 0.000900:   Batch Loss = 0.472827, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5576097369194031, Accu
racy = 0.9127108454704285
Iter #4222976:  Learning rate = 0.000900:   Batch Loss = 0.472966, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5344207286834717, Accu
racy = 0.9224482774734497
Iter #4227072:  Learning rate = 0.000900:   Batch Loss = 0.482201, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5499154329299927, Accu
racy = 0.9257520437240601
Iter #4231168:  Learning rate = 0.000900:   Batch Loss = 0.467960, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5599260926246643, Accu
racy = 0.9193183779716492
Iter #4235264:  Learning rate = 0.000900:   Batch Loss = 0.467116, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.549776017665863, Accur
acy = 0.9181011915206909
Iter #4239360:  Learning rate = 0.000900:   Batch Loss = 0.475838, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5600163340568542, Accu
racy = 0.9154929518699646
Iter #4243456:  Learning rate = 0.000900:   Batch Loss = 0.471235, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5748817324638367, Accu
racy = 0.918796718120575
```

```
Iter #4247552:  Learning rate = 0.000900:   Batch Loss = 0.469328, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5631284117698669, Accu
racy = 0.9224482774734497
Iter #4251648:  Learning rate = 0.000900:   Batch Loss = 0.505420, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5558866262435913, Accu
racy = 0.9146235585212708
Iter #4255744:  Learning rate = 0.000900:   Batch Loss = 0.483889, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5598117709159851, Accu
racy = 0.9128847122192383
Iter #4259840:  Learning rate = 0.000900:   Batch Loss = 0.474291, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5520081520080566, Accu
racy = 0.9137541055679321
Iter #4263936:  Learning rate = 0.000900:   Batch Loss = 0.479487, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5493390560150146, Accu
racy = 0.9200139045715332
Iter #4268032:  Learning rate = 0.000900:   Batch Loss = 0.468505, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5574247241020203, Accu
racy = 0.9172317981719971
Iter #4272128:  Learning rate = 0.000900:   Batch Loss = 0.491057, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5600511431694031, Accu
racy = 0.916536271572113
Iter #4276224:  Learning rate = 0.000900:   Batch Loss = 0.459491, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5835326910018921, Accu
racy = 0.9022778868675232
Iter #4280320:  Learning rate = 0.000900:   Batch Loss = 0.477341, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5516275763511658, Accu
racy = 0.9217527508735657
Iter #4284416:  Learning rate = 0.000900:   Batch Loss = 0.455505, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5500131845474243, Accu
racy = 0.919492244720459
Iter #4288512:  Learning rate = 0.000900:   Batch Loss = 0.474048, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.570388913154602, Accur
acy = 0.9099286794662476
Iter #4292608:  Learning rate = 0.000900:   Batch Loss = 0.537482, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5533510446548462, Accu
racy = 0.915840744972229
Iter #4296704:  Learning rate = 0.000900:   Batch Loss = 0.477010, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5577226877212524, Accu
racy = 0.9241871237754822
Iter #4300800:  Learning rate = 0.000864:   Batch Loss = 0.455799, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5418434143066406, Accu
racy = 0.924360990524292
Iter #4304896:  Learning rate = 0.000864:   Batch Loss = 0.490865, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5732027292251587, Accu
racy = 0.915840744972229
Iter #4308992:  Learning rate = 0.000864:   Batch Loss = 0.469841, Accurac
```

```
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5492960810661316, Accu
racy = 0.9196661710739136
Iter #4313088:  Learning rate = 0.000864:   Batch Loss = 0.481547, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5482841730117798, Accu
racy = 0.9222744107246399
Iter #4317184:  Learning rate = 0.000864:   Batch Loss = 0.438964, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5494326949119568, Accu
racy = 0.9231438040733337
Iter #4321280:  Learning rate = 0.000864:   Batch Loss = 0.507520, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5405603647232056, Accu
racy = 0.923665463924408
Iter #4325376:  Learning rate = 0.000864:   Batch Loss = 0.435824, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5295476913452148, Accu
racy = 0.9266214370727539
Iter #4329472:  Learning rate = 0.000864:   Batch Loss = 0.514426, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5568839311599731, Accu
racy = 0.9219266176223755
Iter #4333568:  Learning rate = 0.000864:   Batch Loss = 0.507056, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.559921145439148, Accur
acy = 0.9127108454704285
Iter #4337664:  Learning rate = 0.000864:   Batch Loss = 0.477629, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5568485856056213, Accu
racy = 0.9170579314231873
Iter #4341760:  Learning rate = 0.000864:   Batch Loss = 0.473875, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5391548871994019, Accu
racy = 0.9231438040733337
Iter #4345856:  Learning rate = 0.000864:   Batch Loss = 0.503937, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5642836093902588, Accu
racy = 0.9167101383209229
Iter #4349952:  Learning rate = 0.000864:   Batch Loss = 0.452316, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.542680025100708, Accur
acy = 0.9212310910224915
Iter #4354048:  Learning rate = 0.000864:   Batch Loss = 0.443773, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5362288951873779, Accu
racy = 0.9259259104728699
Iter #4358144:  Learning rate = 0.000864:   Batch Loss = 0.488581, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5560083389282227, Accu
racy = 0.9193183779716492
Iter #4362240:  Learning rate = 0.000864:   Batch Loss = 0.463518, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5660765767097473, Accu
racy = 0.9203616976737976
Iter #4366336:  Learning rate = 0.000864:   Batch Loss = 0.501694, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5572161674499512, Accu
racy = 0.9144496321678162
Iter #4370432:  Learning rate = 0.000864:   Batch Loss = 0.466209, Accurac
y = 0.953125
```

```
PERFORMANCE ON TEST SET:            Batch Loss = 0.576198160648346, Accur
acy = 0.9156668186187744
Iter #4374528:  Learning rate = 0.000864:   Batch Loss = 0.461406, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:            Batch Loss = 0.5400720834732056, Accu
racy = 0.9229699373245239
Iter #4378624:  Learning rate = 0.000864:   Batch Loss = 0.487341, Accurac
y = 0.96484375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5385991334915161, Accu
racy = 0.9198400378227234
Iter #4382720:  Learning rate = 0.000864:   Batch Loss = 0.496909, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:            Batch Loss = 0.5533068180084229, Accu
racy = 0.919492244720459
Iter #4386816:  Learning rate = 0.000864:   Batch Loss = 0.470061, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:            Batch Loss = 0.5823906660079956, Accu
racy = 0.9135802388191223
Iter #4390912:  Learning rate = 0.000864:   Batch Loss = 0.485883, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5549051761627197, Accu
racy = 0.9144496321678162
Iter #4395008:  Learning rate = 0.000864:   Batch Loss = 0.498014, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5490691661834717, Accu
racy = 0.915145218372345
Iter #4399104:  Learning rate = 0.000864:   Batch Loss = 0.474570, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5821142196655273, Accu
racy = 0.9147974252700806
Iter #4403200:  Learning rate = 0.000830:   Batch Loss = 0.469015, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:            Batch Loss = 0.548109233379364, Accur
acy = 0.9245348572731018
Iter #4407296:  Learning rate = 0.000830:   Batch Loss = 0.479192, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5505634546279907, Accu
racy = 0.9200139045715332
Iter #4411392:  Learning rate = 0.000830:   Batch Loss = 0.468811, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5382132530212402, Accu
racy = 0.9248826503753662
Iter #4415488:  Learning rate = 0.000830:   Batch Loss = 0.455586, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5798559188842773, Accu
racy = 0.9147974252700806
Iter #4419584:  Learning rate = 0.000830:   Batch Loss = 0.497618, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:            Batch Loss = 0.5659564137458801, Accu
racy = 0.9196661710739136
Iter #4423680:  Learning rate = 0.000830:   Batch Loss = 0.542706, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5456240177154541, Accu
racy = 0.9177534580230713
Iter #4427776:  Learning rate = 0.000830:   Batch Loss = 0.503307, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5490056872367859, Accu
racy = 0.9127108454704285
Iter #4431872:  Learning rate = 0.000830:   Batch Loss = 0.462399, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:            Batch Loss = 0.5476162433624268, Accu
```

```
racy = 0.9234915375709534
Iter #4435968:  Learning rate = 0.000830:   Batch Loss = 0.438090, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5641912817955017, Accu
racy = 0.9186228513717651
Iter #4440064:  Learning rate = 0.000830:   Batch Loss = 0.503878, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5494057536125183, Accu
racy = 0.9207094311714172
Iter #4444160:  Learning rate = 0.000830:   Batch Loss = 0.500298, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.568938136100769, Accur
acy = 0.9073204398155212
Iter #4448256:  Learning rate = 0.000830:   Batch Loss = 0.445726, Accurac
y = 0.962890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.554344654083252, Accur
acy = 0.915840744972229
Iter #4452352:  Learning rate = 0.000830:   Batch Loss = 0.500116, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5438268780708313, Accu
racy = 0.9233176708221436
Iter #4456448:  Learning rate = 0.000830:   Batch Loss = 0.443480, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5415405631065369, Accu
racy = 0.9234915375709534
Iter #4460544:  Learning rate = 0.000830:   Batch Loss = 0.469376, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5382788777351379, Accu
racy = 0.9210572242736816
Iter #4464640:  Learning rate = 0.000830:   Batch Loss = 0.456989, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5342342853546143, Accu
racy = 0.9248826503753662
Iter #4468736:  Learning rate = 0.000830:   Batch Loss = 0.445953, Accurac
y = 0.96484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5503368377685547, Accu
racy = 0.924360990524292
Iter #4472832:  Learning rate = 0.000830:   Batch Loss = 0.446963, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.542749285697937, Accur
acy = 0.9186228513717651
Iter #4476928:  Learning rate = 0.000830:   Batch Loss = 0.454820, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5312042832374573, Accu
racy = 0.9248826503753662
Iter #4481024:  Learning rate = 0.000830:   Batch Loss = 0.446911, Accurac
y = 0.962890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5349898338317871, Accu
racy = 0.9238393306732178
Iter #4485120:  Learning rate = 0.000830:   Batch Loss = 0.446361, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5510067343711853, Accu
racy = 0.9227960109710693
Iter #4489216:  Learning rate = 0.000830:   Batch Loss = 0.438043, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5528084635734558, Accu
racy = 0.9177534580230713
Iter #4493312:  Learning rate = 0.000830:   Batch Loss = 0.419948, Accurac
y = 0.9765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5533972382545471, Accu
racy = 0.9248826503753662
```

```
Iter #4497408:  Learning rate = 0.000830:   Batch Loss = 0.478398, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5499640703201294, Accu
racy = 0.9215788841247559
Iter #4501504:  Learning rate = 0.000796:   Batch Loss = 0.447366, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5413095951080322, Accu
racy = 0.923665463924408
Iter #4505600:  Learning rate = 0.000796:   Batch Loss = 0.462827, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5464247465133667, Accu
racy = 0.9245348572731018
Iter #4509696:  Learning rate = 0.000796:   Batch Loss = 0.443897, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5511200428009033, Accu
racy = 0.9198400378227234
Iter #4513792:  Learning rate = 0.000796:   Batch Loss = 0.471315, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5488075017929077, Accu
racy = 0.9184489846229553
Iter #4517888:  Learning rate = 0.000796:   Batch Loss = 0.450064, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5387152433395386, Accu
racy = 0.9241871237754822
Iter #4521984:  Learning rate = 0.000796:   Batch Loss = 0.450191, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5455361604690552, Accu
racy = 0.9233176708221436
Iter #4526080:  Learning rate = 0.000796:   Batch Loss = 0.447551, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5431615710258484, Accu
racy = 0.9226221442222595
Iter #4530176:  Learning rate = 0.000796:   Batch Loss = 0.437191, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5321491360664368, Accu
racy = 0.9248826503753662
Iter #4534272:  Learning rate = 0.000796:   Batch Loss = 1.362633, Accurac
y = 0.71484375
PERFORMANCE ON TEST SET:             Batch Loss = 1.6188602447509766, Accu
racy = 0.6600590944290161
Iter #4538368:  Learning rate = 0.000796:   Batch Loss = 1.099899, Accurac
y = 0.73828125
PERFORMANCE ON TEST SET:             Batch Loss = 1.1506736278533936, Accu
racy = 0.7287428379058838
Iter #4542464:  Learning rate = 0.000796:   Batch Loss = 1.030045, Accurac
y = 0.7421875
PERFORMANCE ON TEST SET:             Batch Loss = 1.0358723402023315, Accu
racy = 0.7254390716552734
Iter #4546560:  Learning rate = 0.000796:   Batch Loss = 0.994991, Accurac
y = 0.783203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9400484561920166, Accu
racy = 0.7643887996673584
Iter #4550656:  Learning rate = 0.000796:   Batch Loss = 0.996321, Accurac
y = 0.7578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.9156200289726257, Accu
racy = 0.7741262316703796
Iter #4554752:  Learning rate = 0.000796:   Batch Loss = 0.905907, Accurac
y = 0.794921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9263104200363159, Accu
racy = 0.7734307050704956
Iter #4558848:  Learning rate = 0.000796:   Batch Loss = 0.954929, Accurac
```

```
y = 0.78125
PERFORMANCE ON TEST SET:             Batch Loss = 0.8991730213165283, Accu
racy = 0.7788210511207581
Iter #4562944:  Learning rate = 0.000796:   Batch Loss = 0.948262, Accurac
y = 0.763671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9072445034980774, Accu
racy = 0.7715179920196533
Iter #4567040:  Learning rate = 0.000796:   Batch Loss = 0.922762, Accurac
y = 0.8046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.9098650217056274, Accu
racy = 0.7709963321685791
Iter #4571136:  Learning rate = 0.000796:   Batch Loss = 0.843883, Accurac
y = 0.810546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.8954111337661743, Accu
racy = 0.781603217124939
Iter #4575232:  Learning rate = 0.000796:   Batch Loss = 0.866315, Accurac
y = 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8856645226478577, Accu
racy = 0.7758650779724121
Iter #4579328:  Learning rate = 0.000796:   Batch Loss = 0.858496, Accurac
y = 0.798828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.8826955556869507, Accu
racy = 0.7836897969245911
Iter #4583424:  Learning rate = 0.000796:   Batch Loss = 0.891273, Accurac
y = 0.771484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.886574387550354, Accur
acy = 0.7878629565238953
Iter #4587520:  Learning rate = 0.000796:   Batch Loss = 0.921406, Accurac
y = 0.779296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.8555097579956055, Accu
racy = 0.7965571284294128
Iter #4591616:  Learning rate = 0.000796:   Batch Loss = 0.868357, Accurac
y = 0.79296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.8606282472610474, Accu
racy = 0.7955138087272644
Iter #4595712:  Learning rate = 0.000796:   Batch Loss = 0.823871, Accurac
y = 0.80859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8292033672332764, Accu
racy = 0.8036863207817078
Iter #4599808:  Learning rate = 0.000796:   Batch Loss = 0.838623, Accurac
y = 0.822265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.8319334387779236, Accu
racy = 0.803512454032898
Iter #4603904:  Learning rate = 0.000765:   Batch Loss = 0.815590, Accurac
y = 0.828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.8276984691619873, Accu
racy = 0.8054251670837402
Iter #4608000:  Learning rate = 0.000765:   Batch Loss = 0.802195, Accurac
y = 0.826171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.8275986313819885, Accu
racy = 0.7976003885269165
Iter #4612096:  Learning rate = 0.000765:   Batch Loss = 0.807894, Accurac
y = 0.8359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8186365365982056, Accu
racy = 0.8043818473815918
Iter #4616192:  Learning rate = 0.000765:   Batch Loss = 0.737761, Accurac
y = 0.853515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.8145201206207275, Accu
racy = 0.8113371729850769
Iter #4620288:  Learning rate = 0.000765:   Batch Loss = 0.782507, Accurac
y = 0.830078125
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.810928225517273, Accur
acy = 0.8073378801345825
Iter #4624384:  Learning rate = 0.000765:   Batch Loss = 0.755398, Accurac
y = 0.8515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.8225899934768677, Accu
racy = 0.8073378801345825
Iter #4628480:  Learning rate = 0.000765:   Batch Loss = 0.769866, Accurac
y = 0.830078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.8451434373855591, Accu
racy = 0.8095983266830444
Iter #4632576:  Learning rate = 0.000765:   Batch Loss = 0.817635, Accurac
y = 0.80859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8131140470504761, Accu
racy = 0.8122065663337708
Iter #4636672:  Learning rate = 0.000765:   Batch Loss = 0.739945, Accurac
y = 0.841796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.795320987701416, Accur
acy = 0.8155103325843811
Iter #4640768:  Learning rate = 0.000765:   Batch Loss = 0.808631, Accurac
y = 0.83203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.8348578810691833, Accu
racy = 0.8132498860359192
Iter #4644864:  Learning rate = 0.000765:   Batch Loss = 0.720791, Accurac
y = 0.86328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7979445457458496, Accu
racy = 0.8122065663337708
Iter #4648960:  Learning rate = 0.000765:   Batch Loss = 0.772445, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.794443666934967, Accur
acy = 0.8170753121376038
Iter #4653056:  Learning rate = 0.000765:   Batch Loss = 0.725579, Accurac
y = 0.83984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7886669635772705, Accu
racy = 0.8149886727333069
Iter #4657152:  Learning rate = 0.000765:   Batch Loss = 0.709997, Accurac
y = 0.86328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.795585036277771, Accur
acy = 0.8111632466316223
Iter #4661248:  Learning rate = 0.000765:   Batch Loss = 0.739750, Accurac
y = 0.85546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7844550609588623, Accu
racy = 0.8132498860359192
Iter #4665344:  Learning rate = 0.000765:   Batch Loss = 0.784835, Accurac
y = 0.83984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7797808647155762, Accu
racy = 0.8177708387374878
Iter #4669440:  Learning rate = 0.000765:   Batch Loss = 0.706598, Accurac
y = 0.857421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7863715887069702, Accu
racy = 0.8198574185371399
Iter #4673536:  Learning rate = 0.000765:   Batch Loss = 0.744913, Accurac
y = 0.8359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7941055297851562, Accu
racy = 0.8198574185371399
Iter #4677632:  Learning rate = 0.000765:   Batch Loss = 0.805680, Accurac
y = 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8273164629936218, Accu
racy = 0.8073378801345825
Iter #4681728:  Learning rate = 0.000765:   Batch Loss = 0.767384, Accurac
y = 0.845703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7839395999908447, Accu
```

```
racy = 0.8191618919372559
Iter #4685824:  Learning rate = 0.000765:   Batch Loss = 0.785818, Accurac
y = 0.822265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7805670499801636, Accu
racy = 0.8163797855377197
Iter #4689920:  Learning rate = 0.000765:   Batch Loss = 0.729580, Accurac
y = 0.8359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7794637680053711, Accu
racy = 0.822639524936676
Iter #4694016:  Learning rate = 0.000765:   Batch Loss = 0.741926, Accurac
y = 0.853515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7748628854751587, Accu
racy = 0.8240305781364441
Iter #4698112:  Learning rate = 0.000765:   Batch Loss = 0.778430, Accurac
y = 0.826171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7982355356216431, Accu
racy = 0.8221178650856018
Iter #4702208:  Learning rate = 0.000734:   Batch Loss = 0.774763, Accurac
y = 0.818359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7883341312408447, Accu
racy = 0.8202051520347595
Iter #4706304:  Learning rate = 0.000734:   Batch Loss = 0.739119, Accurac
y = 0.849609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7707173824310303, Accu
racy = 0.8250738978385925
Iter #4710400:  Learning rate = 0.000734:   Batch Loss = 0.767631, Accurac
y = 0.83984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7721487283706665, Accu
racy = 0.8228134512901306
Iter #4714496:  Learning rate = 0.000734:   Batch Loss = 0.728932, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7888742685317993, Accu
racy = 0.8202051520347595
Iter #4718592:  Learning rate = 0.000734:   Batch Loss = 0.741437, Accurac
y = 0.8359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7747389674186707, Accu
racy = 0.8240305781364441
Iter #4722688:  Learning rate = 0.000734:   Batch Loss = 0.715533, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7704320549964905, Accu
racy = 0.8245522379875183
Iter #4726784:  Learning rate = 0.000734:   Batch Loss = 0.684764, Accurac
y = 0.8671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7720565795898438, Accu
racy = 0.821943998336792
Iter #4730880:  Learning rate = 0.000734:   Batch Loss = 0.766628, Accurac
y = 0.828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7768455743789673, Accu
racy = 0.827508270740509
Iter #4734976:  Learning rate = 0.000734:   Batch Loss = 0.748268, Accurac
y = 0.859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7725459933280945, Accu
racy = 0.822639524936676
Iter #4739072:  Learning rate = 0.000734:   Batch Loss = 0.803308, Accurac
y = 0.822265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7675148844718933, Accu
racy = 0.8254216909408569
Iter #4743168:  Learning rate = 0.000734:   Batch Loss = 0.697662, Accurac
y = 0.8515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7641931772232056, Accu
racy = 0.826812744140625
```

```
Iter #4747264:  Learning rate = 0.000734:   Batch Loss = 0.721783, Accurac
y = 0.857421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7756916284561157, Accu
racy = 0.822639524936676
Iter #4751360:  Learning rate = 0.000734:   Batch Loss = 0.716530, Accurac
y = 0.865234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8062754273414612, Accu
racy = 0.8210746049880981
Iter #4755456:  Learning rate = 0.000734:   Batch Loss = 0.756326, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7935547232627869, Accu
racy = 0.8160319924354553
Iter #4759552:  Learning rate = 0.000734:   Batch Loss = 0.766148, Accurac
y = 0.83984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.8008389472961426, Accu
racy = 0.8193357586860657
Iter #4763648:  Learning rate = 0.000734:   Batch Loss = 0.712250, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7630364894866943, Accu
racy = 0.8250738978385925
Iter #4767744:  Learning rate = 0.000734:   Batch Loss = 0.764291, Accurac
y = 0.837890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7802333831787109, Accu
racy = 0.8280299305915833
Iter #4771840:  Learning rate = 0.000734:   Batch Loss = 0.692447, Accurac
y = 0.849609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.768783688545227, Accur
acy = 0.821943998336792
Iter #4775936:  Learning rate = 0.000734:   Batch Loss = 0.762078, Accurac
y = 0.833984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7674788236618042, Accu
racy = 0.8231611847877502
Iter #4780032:  Learning rate = 0.000734:   Batch Loss = 0.647502, Accurac
y = 0.8828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7614582180976868, Accu
racy = 0.8287254571914673
Iter #4784128:  Learning rate = 0.000734:   Batch Loss = 0.678579, Accurac
y = 0.86328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.761905312538147, Accur
acy = 0.8316814303398132
Iter #4788224:  Learning rate = 0.000734:   Batch Loss = 0.735492, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7764040231704712, Accu
racy = 0.822639524936676
Iter #4792320:  Learning rate = 0.000734:   Batch Loss = 0.734019, Accurac
y = 0.8359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7747718095779419, Accu
racy = 0.821248471736908
Iter #4796416:  Learning rate = 0.000734:   Batch Loss = 0.678195, Accurac
y = 0.873046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7701963186264038, Accu
racy = 0.8245522379875183
Iter #4800512:  Learning rate = 0.000705:   Batch Loss = 0.703270, Accurac
y = 0.8515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.807052731513977, Accur
acy = 0.8236828446388245
Iter #4804608:  Learning rate = 0.000705:   Batch Loss = 0.739412, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7669789791107178, Accu
racy = 0.822639524936676
Iter #4808704:  Learning rate = 0.000705:   Batch Loss = 0.770345, Accurac
```

```
y = 0.82421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7636939287185669, Accu
racy = 0.8243783712387085
Iter #4812800:  Learning rate = 0.000705:   Batch Loss = 0.637461, Accurac
y = 0.89453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7510630488395691, Accu
racy = 0.8339419364929199
Iter #4816896:  Learning rate = 0.000705:   Batch Loss = 0.717419, Accurac
y = 0.859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7559281587600708, Accu
racy = 0.8299426436424255
Iter #4820992:  Learning rate = 0.000705:   Batch Loss = 0.738705, Accurac
y = 0.84765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7563807368278503, Accu
racy = 0.8299426436424255
Iter #4825088:  Learning rate = 0.000705:   Batch Loss = 0.697093, Accurac
y = 0.8671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7409878969192505, Accu
racy = 0.8341158032417297
Iter #4829184:  Learning rate = 0.000705:   Batch Loss = 0.731028, Accurac
y = 0.84765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.764824390411377, Accur
acy = 0.8313336968421936
Iter #4833280:  Learning rate = 0.000705:   Batch Loss = 0.723787, Accurac
y = 0.841796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7532138228416443, Accu
racy = 0.829768717288971
Iter #4837376:  Learning rate = 0.000705:   Batch Loss = 0.728041, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7487767934799194, Accu
racy = 0.835332989692688
Iter #4841472:  Learning rate = 0.000705:   Batch Loss = 0.762303, Accurac
y = 0.833984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7990390062332153, Accu
racy = 0.8292471170425415
Iter #4845568:  Learning rate = 0.000705:   Batch Loss = 0.706289, Accurac
y = 0.8515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.830551028251648, Accur
acy = 0.7998608946800232
Iter #4849664:  Learning rate = 0.000705:   Batch Loss = 0.706056, Accurac
y = 0.861328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.8344505429267883, Accu
racy = 0.8129020929336548
Iter #4853760:  Learning rate = 0.000705:   Batch Loss = 0.763922, Accurac
y = 0.830078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.801252007484436, Accur
acy = 0.8090766668319702
Iter #4857856:  Learning rate = 0.000705:   Batch Loss = 0.710226, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7737032175064087, Accu
racy = 0.8264649510383606
Iter #4861952:  Learning rate = 0.000705:   Batch Loss = 0.708667, Accurac
y = 0.857421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7550537586212158, Accu
racy = 0.8358546495437622
Iter #4866048:  Learning rate = 0.000705:   Batch Loss = 0.736561, Accurac
y = 0.8359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7740389108657837, Accu
racy = 0.821943998336792
Iter #4870144:  Learning rate = 0.000705:   Batch Loss = 0.725581, Accurac
y = 0.833984375
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.771003246307373, Accur
acy = 0.8262910842895508
Iter #4874240:  Learning rate = 0.000705:   Batch Loss = 0.782678, Accurac
y = 0.806640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7525495290756226, Accu
racy = 0.8355068564414978
Iter #4878336:  Learning rate = 0.000705:   Batch Loss = 0.715662, Accurac
y = 0.84375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7490056753158569, Accu
racy = 0.8332464098930359
Iter #4882432:  Learning rate = 0.000705:   Batch Loss = 0.720049, Accurac
y = 0.84765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7781510353088379, Accu
racy = 0.8222917914390564
Iter #4886528:  Learning rate = 0.000705:   Batch Loss = 0.721644, Accurac
y = 0.84765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7502137422561646, Accu
racy = 0.8328986167907715
Iter #4890624:  Learning rate = 0.000705:   Batch Loss = 0.692865, Accurac
y = 0.859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7645775675773621, Accu
racy = 0.8215962648391724
Iter #4894720:  Learning rate = 0.000705:   Batch Loss = 0.745090, Accurac
y = 0.83203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7515039443969727, Accu
racy = 0.8313336968421936
Iter #4898816:  Learning rate = 0.000705:   Batch Loss = 0.668202, Accurac
y = 0.87109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7653660178184509, Accu
racy = 0.8365501761436462
Iter #4902912:  Learning rate = 0.000676:   Batch Loss = 0.669122, Accurac
y = 0.8828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.7448150515556335, Accu
racy = 0.8323769569396973
Iter #4907008:  Learning rate = 0.000676:   Batch Loss = 0.693658, Accurac
y = 0.8671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.737175703048706, Accur
acy = 0.8370718359947205
Iter #4911104:  Learning rate = 0.000676:   Batch Loss = 0.725716, Accurac
y = 0.83984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7374680042266846, Accu
racy = 0.8335941433906555
Iter #4915200:  Learning rate = 0.000676:   Batch Loss = 0.694320, Accurac
y = 0.857421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7711237668991089, Accu
racy = 0.8328986167907715
Iter #4919296:  Learning rate = 0.000676:   Batch Loss = 0.744187, Accurac
y = 0.8359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7676048278808594, Accu
racy = 0.8235089778900146
Iter #4923392:  Learning rate = 0.000676:   Batch Loss = 0.712394, Accurac
y = 0.86328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.762287437915802, Accur
acy = 0.8276821374893188
Iter #4927488:  Learning rate = 0.000676:   Batch Loss = 0.676879, Accurac
y = 0.869140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7436097860336304, Accu
racy = 0.8362023830413818
Iter #4931584:  Learning rate = 0.000676:   Batch Loss = 0.710867, Accurac
y = 0.841796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7558688521385193, Accu
```

```
racy = 0.8283776640892029
Iter #4935680:  Learning rate = 0.000676:   Batch Loss = 0.680738, Accurac
y = 0.8671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7431482076644897, Accu
racy = 0.8315075635910034
Iter #4939776:  Learning rate = 0.000676:   Batch Loss = 0.667892, Accurac
y = 0.873046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7468893527984619, Accu
racy = 0.830464243888855
Iter #4943872:  Learning rate = 0.000676:   Batch Loss = 0.698818, Accurac
y = 0.859375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7345936298370361, Accu
racy = 0.8318553566932678
Iter #4947968:  Learning rate = 0.000676:   Batch Loss = 0.675902, Accurac
y = 0.865234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.741959810256958, Accur
acy = 0.8320292234420776
Iter #4952064:  Learning rate = 0.000676:   Batch Loss = 0.690080, Accurac
y = 0.857421875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7504581212997437, Accu
racy = 0.8313336968421936
Iter #4956160:  Learning rate = 0.000676:   Batch Loss = 0.689968, Accurac
y = 0.853515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7565881013870239, Accu
racy = 0.8316814303398132
Iter #4960256:  Learning rate = 0.000676:   Batch Loss = 0.626494, Accurac
y = 0.87890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7581765651702881, Accu
racy = 0.8308120369911194
Iter #4964352:  Learning rate = 0.000676:   Batch Loss = 0.712502, Accurac
y = 0.853515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7571754455566406, Accu
racy = 0.8292471170425415
Iter #4968448:  Learning rate = 0.000676:   Batch Loss = 0.692451, Accurac
y = 0.865234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7533148527145386, Accu
racy = 0.8252477645874023
Iter #4972544:  Learning rate = 0.000676:   Batch Loss = 0.734628, Accurac
y = 0.833984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.7137759923934937, Accu
racy = 0.8384628891944885
Iter #4976640:  Learning rate = 0.000676:   Batch Loss = 0.646108, Accurac
y = 0.873046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6781765222549438, Accu
racy = 0.8528951406478882
Iter #4980736:  Learning rate = 0.000676:   Batch Loss = 0.631399, Accurac
y = 0.8828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6570994853973389, Accu
racy = 0.8568944334983826
Iter #4984832:  Learning rate = 0.000676:   Batch Loss = 0.610679, Accurac
y = 0.875
PERFORMANCE ON TEST SET:             Batch Loss = 0.64329993724823, Accura
cy = 0.8662841320037842
Iter #4988928:  Learning rate = 0.000676:   Batch Loss = 0.625645, Accurac
y = 0.87109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.653222918510437, Accur
acy = 0.858633279800415
Iter #4993024:  Learning rate = 0.000676:   Batch Loss = 0.636728, Accurac
y = 0.87109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6723655462265015, Accu
racy = 0.8534168004989624
```

```
Iter #4997120:  Learning rate = 0.000676:   Batch Loss = 0.603549, Accurac
y = 0.890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6405497789382935, Accu
racy = 0.8615893125534058
Iter #5001216:  Learning rate = 0.000649:   Batch Loss = 0.606832, Accurac
y = 0.88671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6376426815986633, Accu
racy = 0.8641975522041321
Iter #5005312:  Learning rate = 0.000649:   Batch Loss = 0.631965, Accurac
y = 0.875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6404657363891602, Accu
racy = 0.8648930788040161
Iter #5009408:  Learning rate = 0.000649:   Batch Loss = 0.594215, Accurac
y = 0.890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6216145157814026, Accu
racy = 0.8704572916030884
Iter #5013504:  Learning rate = 0.000649:   Batch Loss = 0.646349, Accurac
y = 0.869140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6646579504013062, Accu
racy = 0.8520257472991943
Iter #5017600:  Learning rate = 0.000649:   Batch Loss = 0.648725, Accurac
y = 0.86328125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6242190599441528, Accu
racy = 0.8692401051521301
Iter #5021696:  Learning rate = 0.000649:   Batch Loss = 0.606588, Accurac
y = 0.89453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6232601404190063, Accu
racy = 0.8681968450546265
Iter #5025792:  Learning rate = 0.000649:   Batch Loss = 0.605022, Accurac
y = 0.88671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6512912511825562, Accu
racy = 0.8574160933494568
Iter #5029888:  Learning rate = 0.000649:   Batch Loss = 0.606436, Accurac
y = 0.884765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.625178337097168, Accur
acy = 0.8695878982543945
Iter #5033984:  Learning rate = 0.000649:   Batch Loss = 0.568243, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6276105046272278, Accu
racy = 0.8661102652549744
Iter #5038080:  Learning rate = 0.000649:   Batch Loss = 0.606493, Accurac
y = 0.865234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6388962864875793, Accu
racy = 0.8641975522041321
Iter #5042176:  Learning rate = 0.000649:   Batch Loss = 0.585359, Accurac
y = 0.888671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6190500259399414, Accu
racy = 0.8666318655014038
Iter #5046272:  Learning rate = 0.000649:   Batch Loss = 0.629144, Accurac
y = 0.87890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6115983724594116, Accu
racy = 0.8794991970062256
Iter #5050368:  Learning rate = 0.000649:   Batch Loss = 0.580067, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6127694845199585, Accu
racy = 0.8683707118034363
Iter #5054464:  Learning rate = 0.000649:   Batch Loss = 0.596904, Accurac
y = 0.8984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6294796466827393, Accu
racy = 0.8654147386550903
Iter #5058560:  Learning rate = 0.000649:   Batch Loss = 0.565656, Accurac
```

```
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6228330135345459, Accu
racy = 0.8721961379051208
Iter #5062656:  Learning rate = 0.000649:   Batch Loss = 0.591481, Accurac
y = 0.888671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6298565864562988, Accu
racy = 0.8685445785522461
Iter #5066752:  Learning rate = 0.000649:   Batch Loss = 0.635907, Accurac
y = 0.85546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6247049570083618, Accu
racy = 0.8694140315055847
Iter #5070848:  Learning rate = 0.000649:   Batch Loss = 0.565872, Accurac
y = 0.900390625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6339026093482971, Accu
racy = 0.8754999041557312
Iter #5074944:  Learning rate = 0.000649:   Batch Loss = 0.579904, Accurac
y = 0.89453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6075023412704468, Accu
racy = 0.8728916645050049
Iter #5079040:  Learning rate = 0.000649:   Batch Loss = 0.571004, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6170530319213867, Accu
racy = 0.8735871911048889
Iter #5083136:  Learning rate = 0.000649:   Batch Loss = 0.532372, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.640191912651062, Accur
acy = 0.8721961379051208
Iter #5087232:  Learning rate = 0.000649:   Batch Loss = 0.957674, Accurac
y = 0.802734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6313338279724121, Accu
racy = 0.8702834248542786
Iter #5091328:  Learning rate = 0.000649:   Batch Loss = 0.615030, Accurac
y = 0.888671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6254103183746338, Accu
racy = 0.876369297504425
Iter #5095424:  Learning rate = 0.000649:   Batch Loss = 0.591550, Accurac
y = 0.892578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6070022583007812, Accu
racy = 0.8831507563591003
Iter #5099520:  Learning rate = 0.000649:   Batch Loss = 0.604830, Accurac
y = 0.869140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6069729328155518, Accu
racy = 0.8704572916030884
Iter #5103616:  Learning rate = 0.000623:   Batch Loss = 0.541776, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5992351770401001, Accu
racy = 0.8767170906066895
Iter #5107712:  Learning rate = 0.000623:   Batch Loss = 0.567367, Accurac
y = 0.908203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6004079580307007, Accu
racy = 0.8840201497077942
Iter #5111808:  Learning rate = 0.000623:   Batch Loss = 0.531595, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6079816222190857, Accu
racy = 0.8735871911048889
Iter #5115904:  Learning rate = 0.000623:   Batch Loss = 0.552543, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6035485863685608, Accu
racy = 0.8831507563591003
Iter #5120000:  Learning rate = 0.000623:   Batch Loss = 0.523166, Accurac
y = 0.927734375
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.6062791347503662, Accu
racy = 0.8831507563591003
Iter #5124096:  Learning rate = 0.000623:   Batch Loss = 0.524104, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6064415574073792, Accu
racy = 0.8836724162101746
Iter #5128192:  Learning rate = 0.000623:   Batch Loss = 0.577382, Accurac
y = 0.892578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5898763537406921, Accu
racy = 0.8909754753112793
Iter #5132288:  Learning rate = 0.000623:   Batch Loss = 0.534861, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5904954075813293, Accu
racy = 0.8920187950134277
Iter #5136384:  Learning rate = 0.000623:   Batch Loss = 0.526029, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5826147794723511, Accu
racy = 0.8935837149620056
Iter #5140480:  Learning rate = 0.000623:   Batch Loss = 0.532618, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5990040302276611, Accu
racy = 0.8888888955116272
Iter #5144576:  Learning rate = 0.000623:   Batch Loss = 0.502356, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5955449342727661, Accu
racy = 0.887671709060669
Iter #5148672:  Learning rate = 0.000623:   Batch Loss = 0.523379, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6032070517539978, Accu
racy = 0.8843679428100586
Iter #5152768:  Learning rate = 0.000623:   Batch Loss = 0.548705, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6002986431121826, Accu
racy = 0.8869761824607849
Iter #5156864:  Learning rate = 0.000623:   Batch Loss = 0.533962, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5867745876312256, Accu
racy = 0.8993218541145325
Iter #5160960:  Learning rate = 0.000623:   Batch Loss = 0.540107, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6005365252494812, Accu
racy = 0.8880195021629333
Iter #5165056:  Learning rate = 0.000623:   Batch Loss = 0.533482, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6084954142570496, Accu
racy = 0.8901060819625854
Iter #5169152:  Learning rate = 0.000623:   Batch Loss = 0.587624, Accurac
y = 0.88671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6107721924781799, Accu
racy = 0.8746305108070374
Iter #5173248:  Learning rate = 0.000623:   Batch Loss = 0.523333, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6050859689712524, Accu
racy = 0.8911493420600891
Iter #5177344:  Learning rate = 0.000623:   Batch Loss = 0.546900, Accurac
y = 0.8984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5897167921066284, Accu
racy = 0.8941053748130798
Iter #5181440:  Learning rate = 0.000623:   Batch Loss = 0.532197, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5866837501525879, Accu
```

```
racy = 0.889062762260437
Iter #5185536:  Learning rate = 0.000623:   Batch Loss = 0.528444, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.585970401763916, Accur
acy = 0.8935837149620056
Iter #5189632:  Learning rate = 0.000623:   Batch Loss = 0.536308, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6001285314559937, Accu
racy = 0.8812380433082581
Iter #5193728:  Learning rate = 0.000623:   Batch Loss = 0.527858, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5996261835098267, Accu
racy = 0.8888888955116272
Iter #5197824:  Learning rate = 0.000623:   Batch Loss = 0.545759, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.589043140411377, Accur
acy = 0.8988001942634583
Iter #5201920:  Learning rate = 0.000599:   Batch Loss = 0.520734, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5986529588699341, Accu
racy = 0.8873239159584045
Iter #5206016:  Learning rate = 0.000599:   Batch Loss = 0.543501, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5733647346496582, Accu
racy = 0.8984524607658386
Iter #5210112:  Learning rate = 0.000599:   Batch Loss = 0.529142, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5901486277580261, Accu
racy = 0.8937575817108154
Iter #5214208:  Learning rate = 0.000599:   Batch Loss = 0.535115, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5839142203330994, Accu
racy = 0.8881933689117432
Iter #5218304:  Learning rate = 0.000599:   Batch Loss = 0.503812, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5854122042655945, Accu
racy = 0.8937575817108154
Iter #5222400:  Learning rate = 0.000599:   Batch Loss = 0.565146, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6144776344299316, Accu
racy = 0.8826290965080261
Iter #5226496:  Learning rate = 0.000599:   Batch Loss = 0.554402, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.581565260887146, Accur
acy = 0.89393150806427
Iter #5230592:  Learning rate = 0.000599:   Batch Loss = 0.517772, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5733064413070679, Accu
racy = 0.8941053748130798
Iter #5234688:  Learning rate = 0.000599:   Batch Loss = 0.498783, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5843336582183838, Accu
racy = 0.8906277418136597
Iter #5238784:  Learning rate = 0.000599:   Batch Loss = 0.522625, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5922949314117432, Accu
racy = 0.8892366290092468
Iter #5242880:  Learning rate = 0.000599:   Batch Loss = 0.523482, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5866338014602661, Accu
racy = 0.8934098482131958
```

```
Iter #5246976:  Learning rate = 0.000599:   Batch Loss = 0.509255, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5753926634788513, Accu
racy = 0.8967136144638062
Iter #5251072:  Learning rate = 0.000599:   Batch Loss = 0.532354, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5872905254364014, Accu
racy = 0.8944531679153442
Iter #5255168:  Learning rate = 0.000599:   Batch Loss = 0.558845, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5921473503112793, Accu
racy = 0.8864545226097107
Iter #5259264:  Learning rate = 0.000599:   Batch Loss = 0.542633, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5927588939666748, Accu
racy = 0.8909754753112793
Iter #5263360:  Learning rate = 0.000599:   Batch Loss = 0.508238, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5697870850563049, Accu
racy = 0.9048861265182495
Iter #5267456:  Learning rate = 0.000599:   Batch Loss = 0.508502, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5748957395553589, Accu
racy = 0.8963658213615417
Iter #5271552:  Learning rate = 0.000599:   Batch Loss = 0.485664, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5802172422409058, Accu
racy = 0.8960180878639221
Iter #5275648:  Learning rate = 0.000599:   Batch Loss = 0.480977, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5714541673660278, Accu
racy = 0.9040166735649109
Iter #5279744:  Learning rate = 0.000599:   Batch Loss = 0.490416, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5813620090484619, Accu
racy = 0.8948009014129639
Iter #5283840:  Learning rate = 0.000599:   Batch Loss = 0.532576, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5735437870025635, Accu
racy = 0.9005390405654907
Iter #5287936:  Learning rate = 0.000599:   Batch Loss = 0.502949, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5909090042114258, Accu
racy = 0.8944531679153442
Iter #5292032:  Learning rate = 0.000599:   Batch Loss = 0.519289, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5690436363220215, Accu
racy = 0.8961919546127319
Iter #5296128:  Learning rate = 0.000599:   Batch Loss = 0.519958, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5835995674133301, Accu
racy = 0.8934098482131958
Iter #5300224:  Learning rate = 0.000575:   Batch Loss = 0.575753, Accurac
y = 0.89453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.614142894744873, Accur
acy = 0.8694140315055847
Iter #5304320:  Learning rate = 0.000575:   Batch Loss = 0.523899, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5811773538589478, Accu
racy = 0.8895844221115112
Iter #5308416:  Learning rate = 0.000575:   Batch Loss = 0.497269, Accurac
```

```
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5779078006744385, Accu
racy = 0.893235981464386
Iter #5312512:  Learning rate = 0.000575:   Batch Loss = 0.513170, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5929833650588989, Accu
racy = 0.8868023157119751
Iter #5316608:  Learning rate = 0.000575:   Batch Loss = 0.531061, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5775030851364136, Accu
racy = 0.89393150806427
Iter #5320704:  Learning rate = 0.000575:   Batch Loss = 0.534269, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5989322662353516, Accu
racy = 0.8847156763076782
Iter #5324800:  Learning rate = 0.000575:   Batch Loss = 0.480212, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5753719806671143, Accu
racy = 0.8996696472167969
Iter #5328896:  Learning rate = 0.000575:   Batch Loss = 0.538044, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5957660675048828, Accu
racy = 0.8923665285110474
Iter #5332992:  Learning rate = 0.000575:   Batch Loss = 0.521069, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5719211101531982, Accu
racy = 0.8979308009147644
Iter #5337088:  Learning rate = 0.000575:   Batch Loss = 0.464261, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5823440551757812, Accu
racy = 0.894627034664154
Iter #5341184:  Learning rate = 0.000575:   Batch Loss = 0.541890, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5619089603424072, Accu
racy = 0.9095809459686279
Iter #5345280:  Learning rate = 0.000575:   Batch Loss = 0.475425, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.56986004114151, Accura
cy = 0.8988001942634583
Iter #5349376:  Learning rate = 0.000575:   Batch Loss = 0.498448, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5769040584564209, Accu
racy = 0.8953225612640381
Iter #5353472:  Learning rate = 0.000575:   Batch Loss = 0.488459, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5720326900482178, Accu
racy = 0.9036689400672913
Iter #5357568:  Learning rate = 0.000575:   Batch Loss = 0.516707, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.578923761844635, Accur
acy = 0.8928881883621216
Iter #5361664:  Learning rate = 0.000575:   Batch Loss = 0.507414, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5837321281433105, Accu
racy = 0.8899322152137756
Iter #5365760:  Learning rate = 0.000575:   Batch Loss = 0.532390, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.597311794757843, Accur
acy = 0.8921926617622375
Iter #5369856:  Learning rate = 0.000575:   Batch Loss = 0.532210, Accurac
y = 0.91796875
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.5668327212333679, Accu
racy = 0.903147280216217
Iter #5373952:  Learning rate = 0.000575:   Batch Loss = 0.504234, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5757237672805786, Accu
racy = 0.8970614075660706
Iter #5378048:  Learning rate = 0.000575:   Batch Loss = 0.525085, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5793222784996033, Accu
racy = 0.8974091410636902
Iter #5382144:  Learning rate = 0.000575:   Batch Loss = 0.478811, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.564099907875061, Accur
acy = 0.901756227016449
Iter #5386240:  Learning rate = 0.000575:   Batch Loss = 0.501696, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5595512390136719, Accu
racy = 0.9029734134674072
Iter #5390336:  Learning rate = 0.000575:   Batch Loss = 0.493645, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5733397006988525, Accu
racy = 0.9001912474632263
Iter #5394432:  Learning rate = 0.000575:   Batch Loss = 0.482017, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5723374485969543, Accu
racy = 0.8986263275146484
Iter #5398528:  Learning rate = 0.000575:   Batch Loss = 0.516552, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5872616767883301, Accu
racy = 0.8984524607658386
Iter #5402624:  Learning rate = 0.000552:   Batch Loss = 0.520270, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5693100690841675, Accu
racy = 0.9008867740631104
Iter #5406720:  Learning rate = 0.000552:   Batch Loss = 0.522332, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.579056978225708, Accur
acy = 0.8986263275146484
Iter #5410816:  Learning rate = 0.000552:   Batch Loss = 0.500091, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5682991743087769, Accu
racy = 0.898278534412384
Iter #5414912:  Learning rate = 0.000552:   Batch Loss = 0.500472, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.559341311454773, Accur
acy = 0.8989741206169128
Iter #5419008:  Learning rate = 0.000552:   Batch Loss = 0.522810, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5652114152908325, Accu
racy = 0.9033211469650269
Iter #5423104:  Learning rate = 0.000552:   Batch Loss = 0.505848, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5734562873840332, Accu
racy = 0.8994957208633423
Iter #5427200:  Learning rate = 0.000552:   Batch Loss = 0.519200, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5626345872879028, Accu
racy = 0.9040166735649109
Iter #5431296:  Learning rate = 0.000552:   Batch Loss = 0.527702, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.562602162361145, Accur
```

```
acy = 0.903147280216217
Iter #5435392:  Learning rate = 0.000552:   Batch Loss = 0.524364, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5694724321365356, Accu
racy = 0.9012345671653748
Iter #5439488:  Learning rate = 0.000552:   Batch Loss = 0.505806, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5689002871513367, Accu
racy = 0.9000173807144165
Iter #5443584:  Learning rate = 0.000552:   Batch Loss = 0.539519, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5804420709609985, Accu
racy = 0.896887481212616
Iter #5447680:  Learning rate = 0.000552:   Batch Loss = 0.526743, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5975028276443481, Accu
racy = 0.892540454864502
Iter #5451776:  Learning rate = 0.000552:   Batch Loss = 0.546216, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5776431560516357, Accu
racy = 0.8843679428100586
Iter #5455872:  Learning rate = 0.000552:   Batch Loss = 0.528407, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5799771547317505, Accu
racy = 0.8951486945152283
Iter #5459968:  Learning rate = 0.000552:   Batch Loss = 0.512810, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5861669778823853, Accu
racy = 0.8911493420600891
Iter #5464064:  Learning rate = 0.000552:   Batch Loss = 0.500400, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.571872889995575, Accur
acy = 0.8974091410636902
Iter #5468160:  Learning rate = 0.000552:   Batch Loss = 0.520896, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5561032891273499, Accu
racy = 0.9052338600158691
Iter #5472256:  Learning rate = 0.000552:   Batch Loss = 0.502796, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5651088953018188, Accu
racy = 0.8979308009147644
Iter #5476352:  Learning rate = 0.000552:   Batch Loss = 0.524036, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5640323758125305, Accu
racy = 0.9010607004165649
Iter #5480448:  Learning rate = 0.000552:   Batch Loss = 0.509071, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5733929872512817, Accu
racy = 0.8967136144638062
Iter #5484544:  Learning rate = 0.000552:   Batch Loss = 0.505769, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5547277331352234, Accu
racy = 0.9048861265182495
Iter #5488640:  Learning rate = 0.000552:   Batch Loss = 0.499469, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.561468243598938, Accur
acy = 0.9027994871139526
Iter #5492736:  Learning rate = 0.000552:   Batch Loss = 0.500994, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.561892032623291, Accur
acy = 0.9047122001647949
```

```
Iter #5496832:  Learning rate = 0.000552:   Batch Loss = 0.529293, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5663996934890747, Accu
racy = 0.903147280216217
Iter #5500928:  Learning rate = 0.000530:   Batch Loss = 0.500072, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5806994438171387, Accu
racy = 0.887671709060669
Iter #5505024:  Learning rate = 0.000530:   Batch Loss = 0.502024, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5700863599777222, Accu
racy = 0.8981046676635742
Iter #5509120:  Learning rate = 0.000530:   Batch Loss = 0.519961, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5899332761764526, Accu
racy = 0.8951486945152283
Iter #5513216:  Learning rate = 0.000530:   Batch Loss = 0.479516, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5682530999183655, Accu
racy = 0.8979308009147644
Iter #5517312:  Learning rate = 0.000530:   Batch Loss = 0.506707, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.56577068567276, Accura
cy = 0.9036689400672913
Iter #5521408:  Learning rate = 0.000530:   Batch Loss = 0.489104, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5605582594871521, Accu
racy = 0.9059293866157532
Iter #5525504:  Learning rate = 0.000530:   Batch Loss = 0.515064, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5687470436096191, Accu
racy = 0.9003651738166809
Iter #5529600:  Learning rate = 0.000530:   Batch Loss = 0.482671, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5530476570129395, Accu
racy = 0.9014084339141846
Iter #5533696:  Learning rate = 0.000530:   Batch Loss = 0.466919, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5529534220695496, Accu
racy = 0.9036689400672913
Iter #5537792:  Learning rate = 0.000530:   Batch Loss = 0.528848, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5575907230377197, Accu
racy = 0.9048861265182495
Iter #5541888:  Learning rate = 0.000530:   Batch Loss = 0.502170, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5718210935592651, Accu
racy = 0.8904538154602051
Iter #5545984:  Learning rate = 0.000530:   Batch Loss = 0.486851, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5489524602890015, Accu
racy = 0.9071465730667114
Iter #5550080:  Learning rate = 0.000530:   Batch Loss = 0.482271, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5525904893875122, Accu
racy = 0.90801602602005
Iter #5554176:  Learning rate = 0.000530:   Batch Loss = 0.496509, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5463560819625854, Accu
racy = 0.9057555198669434
Iter #5558272:  Learning rate = 0.000530:   Batch Loss = 0.528071, Accurac
```

```
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5597552061080933, Accu
racy = 0.9034950733184814
Iter #5562368:  Learning rate = 0.000530:   Batch Loss = 0.490230, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5608291625976562, Accu
racy = 0.9005390405654907
Iter #5566464:  Learning rate = 0.000530:   Batch Loss = 0.465628, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5424004793167114, Accu
racy = 0.9066249132156372
Iter #5570560:  Learning rate = 0.000530:   Batch Loss = 0.456937, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5423455238342285, Accu
racy = 0.9092331528663635
Iter #5574656:  Learning rate = 0.000530:   Batch Loss = 0.463606, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5656065940856934, Accu
racy = 0.8944531679153442
Iter #5578752:  Learning rate = 0.000530:   Batch Loss = 0.458935, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5401479005813599, Accu
racy = 0.9127108454704285
Iter #5582848:  Learning rate = 0.000530:   Batch Loss = 0.457416, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5281507968902588, Accu
racy = 0.9189705848693848
Iter #5586944:  Learning rate = 0.000530:   Batch Loss = 0.505098, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5333746671676636, Accu
racy = 0.915840744972229
Iter #5591040:  Learning rate = 0.000530:   Batch Loss = 0.441274, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5453368425369263, Accu
racy = 0.9127108454704285
Iter #5595136:  Learning rate = 0.000530:   Batch Loss = 0.474573, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5368279814720154, Accu
racy = 0.9184489846229553
Iter #5599232:  Learning rate = 0.000530:   Batch Loss = 0.442949, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5251089334487915, Accu
racy = 0.9184489846229553
Iter #5603328:  Learning rate = 0.000508:   Batch Loss = 0.450859, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5375866889953613, Accu
racy = 0.916536271572113
Iter #5607424:  Learning rate = 0.000508:   Batch Loss = 0.496295, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5176138877868652, Accu
racy = 0.9234915375709534
Iter #5611520:  Learning rate = 0.000508:   Batch Loss = 0.418842, Accurac
y = 0.98046875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5339100360870361, Accu
racy = 0.9170579314231873
Iter #5615616:  Learning rate = 0.000508:   Batch Loss = 0.455964, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5594161748886108, Accu
racy = 0.9069727063179016
Iter #5619712:  Learning rate = 0.000508:   Batch Loss = 0.471938, Accurac
y = 0.9375
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.534960150718689, Accur
acy = 0.9146235585212708
Iter #5623808:  Learning rate = 0.000508:   Batch Loss = 0.488714, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5191594362258911, Accu
racy = 0.9196661710739136
Iter #5627904:  Learning rate = 0.000508:   Batch Loss = 0.446268, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5166451334953308, Accu
racy = 0.9214049577713013
Iter #5632000:  Learning rate = 0.000508:   Batch Loss = 0.448284, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.521916389465332, Accur
acy = 0.9212310910224915
Iter #5636096:  Learning rate = 0.000508:   Batch Loss = 0.464530, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5145135521888733, Accu
racy = 0.9224482774734497
Iter #5640192:  Learning rate = 0.000508:   Batch Loss = 0.482835, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5235961675643921, Accu
racy = 0.920187771320343
Iter #5644288:  Learning rate = 0.000508:   Batch Loss = 0.441882, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5164387226104736, Accu
racy = 0.9229699373245239
Iter #5648384:  Learning rate = 0.000508:   Batch Loss = 0.468255, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5345703959465027, Accu
racy = 0.9174056649208069
Iter #5652480:  Learning rate = 0.000508:   Batch Loss = 0.461471, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5155845880508423, Accu
racy = 0.9269692301750183
Iter #5656576:  Learning rate = 0.000508:   Batch Loss = 0.494003, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 2.7253575325012207, Accu
racy = 0.6125891208648682
Iter #5660672:  Learning rate = 0.000508:   Batch Loss = 1.235504, Accurac
y = 0.70703125
PERFORMANCE ON TEST SET:             Batch Loss = 1.2598435878753662, Accu
racy = 0.7170926928520203
Iter #5664768:  Learning rate = 0.000508:   Batch Loss = 0.733216, Accurac
y = 0.84765625
PERFORMANCE ON TEST SET:             Batch Loss = 0.7691504955291748, Accu
racy = 0.8341158032417297
Iter #5668864:  Learning rate = 0.000508:   Batch Loss = 0.654173, Accurac
y = 0.8671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.7428396344184875, Accu
racy = 0.8509824275970459
Iter #5672960:  Learning rate = 0.000508:   Batch Loss = 0.667621, Accurac
y = 0.8671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6871545314788818, Accu
racy = 0.8610676527023315
Iter #5677056:  Learning rate = 0.000508:   Batch Loss = 0.597426, Accurac
y = 0.888671875
PERFORMANCE ON TEST SET:             Batch Loss = 0.649387001991272, Accur
acy = 0.8725439310073853
Iter #5681152:  Learning rate = 0.000508:   Batch Loss = 0.582972, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.638447642326355, Accur
```

```
acy = 0.8748043775558472
Iter #5685248:  Learning rate = 0.000508:   Batch Loss = 0.598066, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6344250440597534, Accu
racy = 0.8751521706581116
Iter #5689344:  Learning rate = 0.000508:   Batch Loss = 0.614544, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6306390762329102, Accu
racy = 0.8668057918548584
Iter #5693440:  Learning rate = 0.000508:   Batch Loss = 0.530307, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6085430383682251, Accu
racy = 0.887671709060669
Iter #5697536:  Learning rate = 0.000508:   Batch Loss = 0.554238, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6126371026039124, Accu
racy = 0.8751521706581116
Iter #5701632:  Learning rate = 0.000488:   Batch Loss = 0.558240, Accurac
y = 0.90625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6019556522369385, Accu
racy = 0.8828029632568359
Iter #5705728:  Learning rate = 0.000488:   Batch Loss = 0.604492, Accurac
y = 0.890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.6078161001205444, Accu
racy = 0.8796731233596802
Iter #5709824:  Learning rate = 0.000488:   Batch Loss = 0.534747, Accurac
y = 0.912109375
PERFORMANCE ON TEST SET:             Batch Loss = 0.6055391430854797, Accu
racy = 0.8808902502059937
Iter #5713920:  Learning rate = 0.000488:   Batch Loss = 0.589022, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5984035730361938, Accu
racy = 0.8817597031593323
Iter #5718016:  Learning rate = 0.000488:   Batch Loss = 0.528054, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5984244346618652, Accu
racy = 0.8840201497077942
Iter #5722112:  Learning rate = 0.000488:   Batch Loss = 0.541734, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5936249494552612, Accu
racy = 0.889758288860321
Iter #5726208:  Learning rate = 0.000488:   Batch Loss = 0.510269, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5901724100112915, Accu
racy = 0.8899322152137756
Iter #5730304:  Learning rate = 0.000488:   Batch Loss = 0.528597, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5920867323875427, Accu
racy = 0.8829768896102905
Iter #5734400:  Learning rate = 0.000488:   Batch Loss = 0.564779, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5840635895729065, Accu
racy = 0.8902799487113953
Iter #5738496:  Learning rate = 0.000488:   Batch Loss = 0.544701, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5831173658370972, Accu
racy = 0.8904538154602051
Iter #5742592:  Learning rate = 0.000488:   Batch Loss = 0.568409, Accurac
y = 0.89453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5776392221450806, Accu
racy = 0.8935837149620056
```

```
Iter #5746688:  Learning rate = 0.000488:   Batch Loss = 0.525012, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5745381116867065, Accu
racy = 0.8914971351623535
Iter #5750784:  Learning rate = 0.000488:   Batch Loss = 0.544484, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5716984272003174, Accu
racy = 0.8942792415618896
Iter #5754880:  Learning rate = 0.000488:   Batch Loss = 0.577533, Accurac
y = 0.896484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5852501392364502, Accu
racy = 0.8885411024093628
Iter #5758976:  Learning rate = 0.000488:   Batch Loss = 0.524598, Accurac
y = 0.908203125
PERFORMANCE ON TEST SET:             Batch Loss = 0.6106406450271606, Accu
racy = 0.8796731233596802
Iter #5763072:  Learning rate = 0.000488:   Batch Loss = 0.499388, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5821048021316528, Accu
racy = 0.8928881883621216
Iter #5767168:  Learning rate = 0.000488:   Batch Loss = 0.551851, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.582452118396759, Accur
acy = 0.8895844221115112
Iter #5771264:  Learning rate = 0.000488:   Batch Loss = 0.531351, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5737994909286499, Accu
racy = 0.8948009014129639
Iter #5775360:  Learning rate = 0.000488:   Batch Loss = 0.529492, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5647854804992676, Accu
racy = 0.9012345671653748
Iter #5779456:  Learning rate = 0.000488:   Batch Loss = 0.523399, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5648508071899414, Accu
racy = 0.8965397477149963
Iter #5783552:  Learning rate = 0.000488:   Batch Loss = 0.506585, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5627535581588745, Accu
racy = 0.898278534412384
Iter #5787648:  Learning rate = 0.000488:   Batch Loss = 0.490189, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5589456558227539, Accu
racy = 0.901756227016449
Iter #5791744:  Learning rate = 0.000488:   Batch Loss = 0.556796, Accurac
y = 0.8984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5630843639373779, Accu
racy = 0.8975830078125
Iter #5795840:  Learning rate = 0.000488:   Batch Loss = 0.506802, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5581214427947998, Accu
racy = 0.9007129073143005
Iter #5799936:  Learning rate = 0.000488:   Batch Loss = 0.473689, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5606911778450012, Accu
racy = 0.903147280216217
Iter #5804032:  Learning rate = 0.000468:   Batch Loss = 0.554981, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5574866533279419, Accu
racy = 0.9014084339141846
Iter #5808128:  Learning rate = 0.000468:   Batch Loss = 0.499357, Accurac
```

```
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5558179616928101, Accu
racy = 0.9026256203651428
Iter #5812224:  Learning rate = 0.000468:   Batch Loss = 0.542700, Accurac
y = 0.91015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5498577952384949, Accu
racy = 0.9048861265182495
Iter #5816320:  Learning rate = 0.000468:   Batch Loss = 0.485768, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5585446357727051, Accu
racy = 0.9010607004165649
Iter #5820416:  Learning rate = 0.000468:   Batch Loss = 0.456613, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5761573314666748, Accu
racy = 0.8909754753112793
Iter #5824512:  Learning rate = 0.000468:   Batch Loss = 0.551098, Accurac
y = 0.89453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5643998384475708, Accu
racy = 0.8977569341659546
Iter #5828608:  Learning rate = 0.000468:   Batch Loss = 0.535882, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5568498373031616, Accu
racy = 0.8994957208633423
Iter #5832704:  Learning rate = 0.000468:   Batch Loss = 0.510924, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5488922595977783, Accu
racy = 0.8996696472167969
Iter #5836800:  Learning rate = 0.000468:   Batch Loss = 0.502576, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5624866485595703, Accu
racy = 0.9021039605140686
Iter #5840896:  Learning rate = 0.000468:   Batch Loss = 0.583258, Accurac
y = 0.892578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.560107946395874, Accur
acy = 0.8974091410636902
Iter #5844992:  Learning rate = 0.000468:   Batch Loss = 0.516567, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5554372072219849, Accu
racy = 0.9015823602676392
Iter #5849088:  Learning rate = 0.000468:   Batch Loss = 0.536434, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5669137835502625, Accu
racy = 0.9045383334159851
Iter #5853184:  Learning rate = 0.000468:   Batch Loss = 0.512675, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5642839670181274, Accu
racy = 0.8961919546127319
Iter #5857280:  Learning rate = 0.000468:   Batch Loss = 0.532016, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.551220715045929, Accur
acy = 0.9040166735649109
Iter #5861376:  Learning rate = 0.000468:   Batch Loss = 0.489537, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.544786810874939, Accur
acy = 0.9033211469650269
Iter #5865472:  Learning rate = 0.000468:   Batch Loss = 0.514447, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5434743165969849, Accu
racy = 0.9033211469650269
Iter #5869568:  Learning rate = 0.000468:   Batch Loss = 0.505634, Accurac
y = 0.927734375
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.5480972528457642, Accu
racy = 0.8993218541145325
Iter #5873664:  Learning rate = 0.000468:   Batch Loss = 0.455765, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5516607165336609, Accu
racy = 0.8998435139656067
Iter #5877760:  Learning rate = 0.000468:   Batch Loss = 0.545617, Accurac
y = 0.90234375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5519078969955444, Accu
racy = 0.9027994871139526
Iter #5881856:  Learning rate = 0.000468:   Batch Loss = 0.571881, Accurac
y = 0.904296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.6018202900886536, Accu
racy = 0.8788036704063416
Iter #5885952:  Learning rate = 0.000468:   Batch Loss = 0.517490, Accurac
y = 0.916015625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5611716508865356, Accu
racy = 0.8944531679153442
Iter #5890048:  Learning rate = 0.000468:   Batch Loss = 0.496725, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5635651350021362, Accu
racy = 0.8991479873657227
Iter #5894144:  Learning rate = 0.000468:   Batch Loss = 0.530158, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5652843713760376, Accu
racy = 0.8981046676635742
Iter #5898240:  Learning rate = 0.000468:   Batch Loss = 0.511501, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5579431056976318, Accu
racy = 0.9005390405654907
Iter #5902336:  Learning rate = 0.000450:   Batch Loss = 0.496229, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5508986711502075, Accu
racy = 0.9094070792198181
Iter #5906432:  Learning rate = 0.000450:   Batch Loss = 0.519414, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5598577857017517, Accu
racy = 0.9001912474632263
Iter #5910528:  Learning rate = 0.000450:   Batch Loss = 0.523986, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5547407865524292, Accu
racy = 0.9022778868675232
Iter #5914624:  Learning rate = 0.000450:   Batch Loss = 0.461821, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5629556179046631, Accu
racy = 0.8977569341659546
Iter #5918720:  Learning rate = 0.000450:   Batch Loss = 0.509557, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5479402542114258, Accu
racy = 0.9050599932670593
Iter #5922816:  Learning rate = 0.000450:   Batch Loss = 0.494503, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5491771697998047, Accu
racy = 0.9019300937652588
Iter #5926912:  Learning rate = 0.000450:   Batch Loss = 0.502618, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5422751307487488, Accu
racy = 0.9061033129692078
Iter #5931008:  Learning rate = 0.000450:   Batch Loss = 0.492690, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5379216074943542, Accu
```

```
racy = 0.9040166735649109
Iter #5935104:  Learning rate = 0.000450:   Batch Loss = 0.458353, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5405788421630859, Accu
racy = 0.9033211469650269
Iter #5939200:  Learning rate = 0.000450:   Batch Loss = 0.499651, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5325419902801514, Accu
racy = 0.9090592861175537
Iter #5943296:  Learning rate = 0.000450:   Batch Loss = 0.484324, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5496894717216492, Accu
racy = 0.9029734134674072
Iter #5947392:  Learning rate = 0.000450:   Batch Loss = 0.485961, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5368386507034302, Accu
racy = 0.9074943661689758
Iter #5951488:  Learning rate = 0.000450:   Batch Loss = 0.492458, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5363049507141113, Accu
racy = 0.9078420996665955
Iter #5955584:  Learning rate = 0.000450:   Batch Loss = 0.464733, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5331180095672607, Accu
racy = 0.9057555198669434
Iter #5959680:  Learning rate = 0.000450:   Batch Loss = 0.469515, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5202245712280273, Accu
racy = 0.9182750582695007
Iter #5963776:  Learning rate = 0.000450:   Batch Loss = 0.493377, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.534593939781189, Accur
acy = 0.9083637595176697
Iter #5967872:  Learning rate = 0.000450:   Batch Loss = 0.492931, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5390146374702454, Accu
racy = 0.9048861265182495
Iter #5971968:  Learning rate = 0.000450:   Batch Loss = 0.509387, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5522154569625854, Accu
racy = 0.9055816531181335
Iter #5976064:  Learning rate = 0.000450:   Batch Loss = 0.515528, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5422375798225403, Accu
racy = 0.9081898927688599
Iter #5980160:  Learning rate = 0.000450:   Batch Loss = 0.476538, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5507104992866516, Accu
racy = 0.9061033129692078
Iter #5984256:  Learning rate = 0.000450:   Batch Loss = 0.473593, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5337371826171875, Accu
racy = 0.9141018986701965
Iter #5988352:  Learning rate = 0.000450:   Batch Loss = 0.547121, Accurac
y = 0.8984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5502095222473145, Accu
racy = 0.8993218541145325
Iter #5992448:  Learning rate = 0.000450:   Batch Loss = 0.447996, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.530012845993042, Accur
acy = 0.9120153188705444
```

```
Iter #5996544:  Learning rate = 0.000450:   Batch Loss = 0.486094, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5428813099861145, Accu
racy = 0.9041905999183655
Iter #6000640:  Learning rate = 0.000432:   Batch Loss = 0.526553, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5351192951202393, Accu
racy = 0.9144496321678162
Iter #6004736:  Learning rate = 0.000432:   Batch Loss = 0.502216, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5532134771347046, Accu
racy = 0.9008867740631104
Iter #6008832:  Learning rate = 0.000432:   Batch Loss = 0.482652, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5335649847984314, Accu
racy = 0.91166752576828
Iter #6012928:  Learning rate = 0.000432:   Batch Loss = 0.502523, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5359315276145935, Accu
racy = 0.90801602602005
Iter #6017024:  Learning rate = 0.000432:   Batch Loss = 0.499819, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5280667543411255, Accu
racy = 0.9106242656707764
Iter #6021120:  Learning rate = 0.000432:   Batch Loss = 0.508590, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5255612134933472, Accu
racy = 0.9121891856193542
Iter #6025216:  Learning rate = 0.000432:   Batch Loss = 0.490720, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.524354100227356, Accur
acy = 0.9144496321678162
Iter #6029312:  Learning rate = 0.000432:   Batch Loss = 0.498118, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5348780155181885, Accu
racy = 0.9088854193687439
Iter #6033408:  Learning rate = 0.000432:   Batch Loss = 0.471534, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5562498569488525, Accu
racy = 0.9015823602676392
Iter #6037504:  Learning rate = 0.000432:   Batch Loss = 0.478854, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5471363067626953, Accu
racy = 0.9034950733184814
Iter #6041600:  Learning rate = 0.000432:   Batch Loss = 0.520507, Accurac
y = 0.91796875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5368685722351074, Accu
racy = 0.9078420996665955
Iter #6045696:  Learning rate = 0.000432:   Batch Loss = 0.454643, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5477537512779236, Accu
racy = 0.9027994871139526
Iter #6049792:  Learning rate = 0.000432:   Batch Loss = 0.501810, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5304393172264099, Accu
racy = 0.9135802388191223
Iter #6053888:  Learning rate = 0.000432:   Batch Loss = 0.482727, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5285792350769043, Accu
racy = 0.9099286794662476
Iter #6057984:  Learning rate = 0.000432:   Batch Loss = 0.483595, Accurac
```

```
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5215131044387817, Accu
racy = 0.9142757654190063
Iter #6062080:  Learning rate = 0.000432:   Batch Loss = 0.516545, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5309970378875732, Accu
racy = 0.910276472568512
Iter #6066176:  Learning rate = 0.000432:   Batch Loss = 0.496776, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.522383987903595, Accur
acy = 0.915840744972229
Iter #6070272:  Learning rate = 0.000432:   Batch Loss = 0.477282, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5278542041778564, Accu
racy = 0.9111458659172058
Iter #6074368:  Learning rate = 0.000432:   Batch Loss = 0.460933, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5320858955383301, Accu
racy = 0.910276472568512
Iter #6078464:  Learning rate = 0.000432:   Batch Loss = 0.481171, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5349958539009094, Accu
racy = 0.9081898927688599
Iter #6082560:  Learning rate = 0.000432:   Batch Loss = 0.452956, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5284156203269958, Accu
racy = 0.9120153188705444
Iter #6086656:  Learning rate = 0.000432:   Batch Loss = 0.465371, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5249704718589783, Accu
racy = 0.9121891856193542
Iter #6090752:  Learning rate = 0.000432:   Batch Loss = 0.486637, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5251203775405884, Accu
racy = 0.9170579314231873
Iter #6094848:  Learning rate = 0.000432:   Batch Loss = 0.484022, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5229935646057129, Accu
racy = 0.915840744972229
Iter #6098944:  Learning rate = 0.000432:   Batch Loss = 0.473130, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5268774032592773, Accu
racy = 0.9142757654190063
Iter #6103040:  Learning rate = 0.000414:   Batch Loss = 0.471625, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5220902562141418, Accu
racy = 0.9156668186187744
Iter #6107136:  Learning rate = 0.000414:   Batch Loss = 0.507646, Accurac
y = 0.921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5218015909194946, Accu
racy = 0.9128847122192383
Iter #6111232:  Learning rate = 0.000414:   Batch Loss = 0.487504, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5169198513031006, Accu
racy = 0.9141018986701965
Iter #6115328:  Learning rate = 0.000414:   Batch Loss = 0.501818, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5231406688690186, Accu
racy = 0.9130585789680481
Iter #6119424:  Learning rate = 0.000414:   Batch Loss = 0.462825, Accurac
y = 0.94921875
```

```
PERFORMANCE ON TEST SET:            Batch Loss = 0.51998370885849, Accura
cy = 0.9179273247718811
Iter #6123520:  Learning rate = 0.000414:   Batch Loss = 0.454208, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:            Batch Loss = 0.5156670212745667, Accu
racy = 0.9170579314231873
Iter #6127616:  Learning rate = 0.000414:   Batch Loss = 0.474119, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5223578214645386, Accu
racy = 0.9139280319213867
Iter #6131712:  Learning rate = 0.000414:   Batch Loss = 0.458377, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:            Batch Loss = 0.516137421131134, Accur
acy = 0.9198400378227234
Iter #6135808:  Learning rate = 0.000414:   Batch Loss = 0.477677, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:            Batch Loss = 0.5441007614135742, Accu
racy = 0.903147280216217
Iter #6139904:  Learning rate = 0.000414:   Batch Loss = 0.476607, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:            Batch Loss = 0.506463348865509, Accur
acy = 0.9210572242736816
Iter #6144000:  Learning rate = 0.000414:   Batch Loss = 0.512045, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:            Batch Loss = 0.5065855979919434, Accu
racy = 0.9215788841247559
Iter #6148096:  Learning rate = 0.000414:   Batch Loss = 0.511487, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5143939256668091, Accu
racy = 0.9175795316696167
Iter #6152192:  Learning rate = 0.000414:   Batch Loss = 0.462443, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:            Batch Loss = 0.5145276188850403, Accu
racy = 0.9179273247718811
Iter #6156288:  Learning rate = 0.000414:   Batch Loss = 0.485912, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5147134065628052, Accu
racy = 0.918796718120575
Iter #6160384:  Learning rate = 0.000414:   Batch Loss = 0.467867, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:            Batch Loss = 0.5294597148895264, Accu
racy = 0.9114936590194702
Iter #6164480:  Learning rate = 0.000414:   Batch Loss = 0.509855, Accurac
y = 0.927734375
PERFORMANCE ON TEST SET:            Batch Loss = 0.5471370816230774, Accu
racy = 0.9073204398155212
Iter #6168576:  Learning rate = 0.000414:   Batch Loss = 0.478693, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:            Batch Loss = 0.5182875990867615, Accu
racy = 0.9147974252700806
Iter #6172672:  Learning rate = 0.000414:   Batch Loss = 0.508447, Accurac
y = 0.9140625
PERFORMANCE ON TEST SET:            Batch Loss = 0.5208573341369629, Accu
racy = 0.9125369787216187
Iter #6176768:  Learning rate = 0.000414:   Batch Loss = 0.485063, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:            Batch Loss = 0.5356755256652832, Accu
racy = 0.9118413925170898
Iter #6180864:  Learning rate = 0.000414:   Batch Loss = 0.436034, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:            Batch Loss = 0.5330668687820435, Accu
```

```
racy = 0.9111458659172058
Iter #6184960:  Learning rate = 0.000414:   Batch Loss = 0.511590, Accurac
y = 0.919921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5138932466506958, Accu
racy = 0.9217527508735657
Iter #6189056:  Learning rate = 0.000414:   Batch Loss = 0.482040, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.527296781539917, Accur
acy = 0.9088854193687439
Iter #6193152:  Learning rate = 0.000414:   Batch Loss = 0.472952, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5345615744590759, Accu
racy = 0.9142757654190063
Iter #6197248:  Learning rate = 0.000414:   Batch Loss = 0.479489, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5273844599723816, Accu
racy = 0.9104503393173218
Iter #6201344:  Learning rate = 0.000398:   Batch Loss = 0.482353, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5224936008453369, Accu
racy = 0.9132325053215027
Iter #6205440:  Learning rate = 0.000398:   Batch Loss = 0.492613, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5299658179283142, Accu
racy = 0.9099286794662476
Iter #6209536:  Learning rate = 0.000398:   Batch Loss = 0.475144, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5065867900848389, Accu
racy = 0.9198400378227234
Iter #6213632:  Learning rate = 0.000398:   Batch Loss = 0.455280, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5064140558242798, Accu
racy = 0.9203616976737976
Iter #6217728:  Learning rate = 0.000398:   Batch Loss = 0.477373, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5096341371536255, Accu
racy = 0.9198400378227234
Iter #6221824:  Learning rate = 0.000398:   Batch Loss = 0.444098, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5073831081390381, Accu
racy = 0.9226221442222595
Iter #6225920:  Learning rate = 0.000398:   Batch Loss = 0.468651, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.49890583753585815, Acc
uracy = 0.9274908900260925
Iter #6230016:  Learning rate = 0.000398:   Batch Loss = 0.480405, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.517809271812439, Accur
acy = 0.915840744972229
Iter #6234112:  Learning rate = 0.000398:   Batch Loss = 0.510440, Accurac
y = 0.92578125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5194019675254822, Accu
racy = 0.9200139045715332
Iter #6238208:  Learning rate = 0.000398:   Batch Loss = 0.484854, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5189868807792664, Accu
racy = 0.9144496321678162
Iter #6242304:  Learning rate = 0.000398:   Batch Loss = 0.493787, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5121821761131287, Accu
racy = 0.9198400378227234
```

```
Iter #6246400:  Learning rate = 0.000398:   Batch Loss = 0.420383, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.49991297721862793, Acc
uracy = 0.9233176708221436
Iter #6250496:  Learning rate = 0.000398:   Batch Loss = 0.469020, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5042952299118042, Accu
racy = 0.9217527508735657
Iter #6254592:  Learning rate = 0.000398:   Batch Loss = 0.435135, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5008261203765869, Accu
racy = 0.9241871237754822
Iter #6258688:  Learning rate = 0.000398:   Batch Loss = 0.438718, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5022792816162109, Accu
racy = 0.9271430969238281
Iter #6262784:  Learning rate = 0.000398:   Batch Loss = 0.497493, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5110781788825989, Accu
racy = 0.9215788841247559
Iter #6266880:  Learning rate = 0.000398:   Batch Loss = 0.470935, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5089581608772278, Accu
racy = 0.9186228513717651
Iter #6270976:  Learning rate = 0.000398:   Batch Loss = 0.462204, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5002334713935852, Accu
racy = 0.9294036030769348
Iter #6275072:  Learning rate = 0.000398:   Batch Loss = 0.456872, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5019201636314392, Accu
racy = 0.9227960109710693
Iter #6279168:  Learning rate = 0.000398:   Batch Loss = 0.470888, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5010544061660767, Accu
racy = 0.9248826503753662
Iter #6283264:  Learning rate = 0.000398:   Batch Loss = 0.490949, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5175583958625793, Accu
racy = 0.9193183779716492
Iter #6287360:  Learning rate = 0.000398:   Batch Loss = 0.500551, Accurac
y = 0.923828125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5035141110420227, Accu
racy = 0.9207094311714172
Iter #6291456:  Learning rate = 0.000398:   Batch Loss = 0.504762, Accurac
y = 0.931640625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5029176473617554, Accu
racy = 0.923665463924408
Iter #6295552:  Learning rate = 0.000398:   Batch Loss = 0.451777, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5051506757736206, Accu
racy = 0.9215788841247559
Iter #6299648:  Learning rate = 0.000398:   Batch Loss = 0.458114, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.4982941150665283, Accu
racy = 0.9245348572731018
Iter #6303744:  Learning rate = 0.000382:   Batch Loss = 0.449738, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.49341559410095215, Acc
uracy = 0.9314901828765869
Iter #6307840:  Learning rate = 0.000382:   Batch Loss = 0.438272, Accurac
```

```
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.4982990026473999, Accu
racy = 0.9234915375709534
Iter #6311936:  Learning rate = 0.000382:   Batch Loss = 0.437481, Accurac
y = 0.962890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5004578828811646, Accu
racy = 0.924360990524292
Iter #6316032:  Learning rate = 0.000382:   Batch Loss = 0.461175, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5057970285415649, Accu
racy = 0.9266214370727539
Iter #6320128:  Learning rate = 0.000382:   Batch Loss = 0.477684, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5106037855148315, Accu
racy = 0.9248826503753662
Iter #6324224:  Learning rate = 0.000382:   Batch Loss = 0.493746, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5330840349197388, Accu
racy = 0.915145218372345
Iter #6328320:  Learning rate = 0.000382:   Batch Loss = 0.473326, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.5151533484458923, Accu
racy = 0.915840744972229
Iter #6332416:  Learning rate = 0.000382:   Batch Loss = 0.427529, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5024794936180115, Accu
racy = 0.9224482774734497
Iter #6336512:  Learning rate = 0.000382:   Batch Loss = 0.470202, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5059066414833069, Accu
racy = 0.9189705848693848
Iter #6340608:  Learning rate = 0.000382:   Batch Loss = 0.467920, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5104122161865234, Accu
racy = 0.9221004843711853
Iter #6344704:  Learning rate = 0.000382:   Batch Loss = 0.424758, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5062896013259888, Accu
racy = 0.920187771320343
Iter #6348800:  Learning rate = 0.000382:   Batch Loss = 0.440901, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.49443453550338745, Acc
uracy = 0.9274908900260925
Iter #6352896:  Learning rate = 0.000382:   Batch Loss = 0.433629, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.49735450744628906, Acc
uracy = 0.9273169636726379
Iter #6356992:  Learning rate = 0.000382:   Batch Loss = 0.459374, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5136657357215881, Accu
racy = 0.9154929518699646
Iter #6361088:  Learning rate = 0.000382:   Batch Loss = 0.483214, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.49596065282821655, Acc
uracy = 0.9299252033233643
Iter #6365184:  Learning rate = 0.000382:   Batch Loss = 0.432361, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5088725686073303, Accu
racy = 0.918796718120575
Iter #6369280:  Learning rate = 0.000382:   Batch Loss = 0.440131, Accurac
y = 0.947265625
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.49402183294296265, Acc
uracy = 0.9271430969238281
Iter #6373376:  Learning rate = 0.000382:   Batch Loss = 0.470918, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5066940784454346, Accu
racy = 0.9203616976737976
Iter #6377472:  Learning rate = 0.000382:   Batch Loss = 0.447923, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5029664635658264, Accu
racy = 0.9264475703239441
Iter #6381568:  Learning rate = 0.000382:   Batch Loss = 0.445337, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.4974236488342285, Accu
racy = 0.9257520437240601
Iter #6385664:  Learning rate = 0.000382:   Batch Loss = 0.423290, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5016387104988098, Accu
racy = 0.928012490272522
Iter #6389760:  Learning rate = 0.000382:   Batch Loss = 0.420264, Accurac
y = 0.96484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.49423131346702576, Acc
uracy = 0.9273169636726379
Iter #6393856:  Learning rate = 0.000382:   Batch Loss = 0.446665, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4922485947608948, Accu
racy = 0.9292296767234802
Iter #6397952:  Learning rate = 0.000382:   Batch Loss = 0.431712, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.49501436948776245, Acc
uracy = 0.928012490272522
Iter #6402048:  Learning rate = 0.000367:   Batch Loss = 0.445939, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.4904397130012512, Accu
racy = 0.9290558099746704
Iter #6406144:  Learning rate = 0.000367:   Batch Loss = 0.470970, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4905821979045868, Accu
racy = 0.9292296767234802
Iter #6410240:  Learning rate = 0.000367:   Batch Loss = 0.470235, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.49240392446517944, Acc
uracy = 0.9274908900260925
Iter #6414336:  Learning rate = 0.000367:   Batch Loss = 0.439433, Accurac
y = 0.962890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.487740159034729, Accur
acy = 0.9294036030769348
Iter #6418432:  Learning rate = 0.000367:   Batch Loss = 0.440221, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.48263922333717346, Acc
uracy = 0.9300991296768188
Iter #6422528:  Learning rate = 0.000367:   Batch Loss = 0.468261, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.4900127947330475, Accu
racy = 0.9271430969238281
Iter #6426624:  Learning rate = 0.000367:   Batch Loss = 0.422209, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.4920993745326996, Accu
racy = 0.9273169636726379
Iter #6430720:  Learning rate = 0.000367:   Batch Loss = 0.439442, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.4856739640235901, Accu
```

```
racy = 0.9309685230255127
Iter #6434816:  Learning rate = 0.000367:   Batch Loss = 0.441253, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5079723000526428, Accu
racy = 0.9227960109710693
Iter #6438912:  Learning rate = 0.000367:   Batch Loss = 0.471813, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5097109079360962, Accu
racy = 0.9238393306732178
Iter #6443008:  Learning rate = 0.000367:   Batch Loss = 0.416902, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4974249601364136, Accu
racy = 0.9255781769752502
Iter #6447104:  Learning rate = 0.000367:   Batch Loss = 0.456343, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.49597057700157166, Acc
uracy = 0.9248826503753662
Iter #6451200:  Learning rate = 0.000367:   Batch Loss = 0.429893, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4992607831954956, Accu
racy = 0.919492244720459
Iter #6455296:  Learning rate = 0.000367:   Batch Loss = 0.454422, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.4886908531188965, Accu
racy = 0.933576762676239
Iter #6459392:  Learning rate = 0.000367:   Batch Loss = 0.449776, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.4902576208114624, Accu
racy = 0.9304468631744385
Iter #6463488:  Learning rate = 0.000367:   Batch Loss = 0.463891, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.4969994127750397, Accu
racy = 0.9222744107246399
Iter #6467584:  Learning rate = 0.000367:   Batch Loss = 0.479573, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.47934800386428833, Acc
uracy = 0.9398365616798401
Iter #6471680:  Learning rate = 0.000367:   Batch Loss = 0.466246, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.488284707069397, Accur
acy = 0.9269692301750183
Iter #6475776:  Learning rate = 0.000367:   Batch Loss = 0.452187, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.4936690032482147, Accu
racy = 0.9287080764770508
Iter #6479872:  Learning rate = 0.000367:   Batch Loss = 0.429760, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.4855777621269226, Accu
racy = 0.9295774698257446
Iter #6483968:  Learning rate = 0.000367:   Batch Loss = 0.449006, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.4815685749053955, Accu
racy = 0.9325334429740906
Iter #6488064:  Learning rate = 0.000367:   Batch Loss = 0.449325, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.48492181301116943, Acc
uracy = 0.9360111355781555
Iter #6492160:  Learning rate = 0.000367:   Batch Loss = 0.473569, Accurac
y = 0.94140625
PERFORMANCE ON TEST SET:             Batch Loss = 0.4960833787918091, Accu
racy = 0.9260998368263245
```

```
Iter #6496256:  Learning rate = 0.000367:   Batch Loss = 0.438097, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.48223361372947693, Acc
uracy = 0.9353156089782715
Iter #6500352:  Learning rate = 0.000352:   Batch Loss = 0.405896, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.4819631576538086, Accu
racy = 0.9306207895278931
Iter #6504448:  Learning rate = 0.000352:   Batch Loss = 0.489770, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.48764005303382874, Acc
uracy = 0.9307946562767029
Iter #6508544:  Learning rate = 0.000352:   Batch Loss = 0.484129, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.48397183418273926, Acc
uracy = 0.9325334429740906
Iter #6512640:  Learning rate = 0.000352:   Batch Loss = 0.416671, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4787813425064087, Accu
racy = 0.9337506294250488
Iter #6516736:  Learning rate = 0.000352:   Batch Loss = 0.456895, Accurac
y = 0.9375
PERFORMANCE ON TEST SET:             Batch Loss = 0.47906243801116943, Acc
uracy = 0.932185709476471
Iter #6520832:  Learning rate = 0.000352:   Batch Loss = 0.419420, Accurac
y = 0.962890625
PERFORMANCE ON TEST SET:             Batch Loss = 0.48604798316955566, Acc
uracy = 0.9273169636726379
Iter #6524928:  Learning rate = 0.000352:   Batch Loss = 0.457503, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5039175748825073, Accu
racy = 0.9259259104728699
Iter #6529024:  Learning rate = 0.000352:   Batch Loss = 0.444129, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.48757463693618774, Acc
uracy = 0.9273169636726379
Iter #6533120:  Learning rate = 0.000352:   Batch Loss = 0.464601, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.49462607502937317, Acc
uracy = 0.9294036030769348
Iter #6537216:  Learning rate = 0.000352:   Batch Loss = 0.445692, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.48320600390434265, Acc
uracy = 0.9281864166259766
Iter #6541312:  Learning rate = 0.000352:   Batch Loss = 0.415814, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.47964614629745483, Acc
uracy = 0.9330551028251648
Iter #6545408:  Learning rate = 0.000352:   Batch Loss = 0.473722, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.49038946628570557, Acc
uracy = 0.9327073693275452
Iter #6549504:  Learning rate = 0.000352:   Batch Loss = 0.478613, Accurac
y = 0.939453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.48820868134498596, Acc
uracy = 0.9300991296768188
Iter #6553600:  Learning rate = 0.000352:   Batch Loss = 0.420544, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4823552966117859, Accu
racy = 0.9290558099746704
Iter #6557696:  Learning rate = 0.000352:   Batch Loss = 0.426537, Accurac
```

```
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.483961820602417, Accur
acy = 0.9311423897743225
Iter #6561792:  Learning rate = 0.000352:   Batch Loss = 0.457072, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5039494037628174, Accu
racy = 0.9248826503753662
Iter #6565888:  Learning rate = 0.000352:   Batch Loss = 0.458534, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.49152928590774536, Acc
uracy = 0.9288819432258606
Iter #6569984:  Learning rate = 0.000352:   Batch Loss = 0.443734, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.49359166622161865, Acc
uracy = 0.9309685230255127
Iter #6574080:  Learning rate = 0.000352:   Batch Loss = 0.445463, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.47526293992996216, Acc
uracy = 0.9353156089782715
Iter #6578176:  Learning rate = 0.000352:   Batch Loss = 0.433751, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.49067842960357666, Acc
uracy = 0.9254042506217957
Iter #6582272:  Learning rate = 0.000352:   Batch Loss = 0.437683, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.4827089011669159, Accu
racy = 0.9309685230255127
Iter #6586368:  Learning rate = 0.000352:   Batch Loss = 0.418362, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.48163485527038574, Acc
uracy = 0.9297513365745544
Iter #6590464:  Learning rate = 0.000352:   Batch Loss = 0.424748, Accurac
y = 0.96484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.4784623086452484, Accu
racy = 0.934272289276123
Iter #6594560:  Learning rate = 0.000352:   Batch Loss = 0.414966, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.47532474994659424, Acc
uracy = 0.9346200823783875
Iter #6598656:  Learning rate = 0.000352:   Batch Loss = 0.423112, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.4746381640434265, Accu
racy = 0.9353156089782715
Iter #6602752:  Learning rate = 0.000338:   Batch Loss = 0.415477, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4713743329048157, Accu
racy = 0.9375760555267334
Iter #6606848:  Learning rate = 0.000338:   Batch Loss = 0.451693, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.47667232155799866, Acc
uracy = 0.9334028959274292
Iter #6610944:  Learning rate = 0.000338:   Batch Loss = 0.434311, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4766879677772522, Accu
racy = 0.9361850023269653
Iter #6615040:  Learning rate = 0.000338:   Batch Loss = 0.429695, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.488129585981369, Accur
acy = 0.9285341501235962
Iter #6619136:  Learning rate = 0.000338:   Batch Loss = 0.414107, Accurac
y = 0.953125
```

```
PERFORMANCE ON TEST SET:             Batch Loss = 0.47235316038131714, Acc
uracy = 0.9367066621780396
Iter #6623232:  Learning rate = 0.000338:   Batch Loss = 0.463162, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.493997186422348, Accur
acy = 0.9262737035751343
Iter #6627328:  Learning rate = 0.000338:   Batch Loss = 0.435121, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.47775202989578247, Acc
uracy = 0.9307946562767029
Iter #6631424:  Learning rate = 0.000338:   Batch Loss = 0.420820, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.47904324531555176, Acc
uracy = 0.932881236076355
Iter #6635520:  Learning rate = 0.000338:   Batch Loss = 0.414482, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.4913311004638672, Accu
racy = 0.9264475703239441
Iter #6639616:  Learning rate = 0.000338:   Batch Loss = 0.435272, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.502052903175354, Accur
acy = 0.925056517124176
Iter #6643712:  Learning rate = 0.000338:   Batch Loss = 0.506858, Accurac
y = 0.9296875
PERFORMANCE ON TEST SET:             Batch Loss = 0.5130656361579895, Accu
racy = 0.9174056649208069
Iter #6647808:  Learning rate = 0.000338:   Batch Loss = 0.466133, Accurac
y = 0.93359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.5024889707565308, Accu
racy = 0.9266214370727539
Iter #6651904:  Learning rate = 0.000338:   Batch Loss = 0.426875, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.48968350887298584, Acc
uracy = 0.9292296767234802
Iter #6656000:  Learning rate = 0.000338:   Batch Loss = 0.454950, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.5063726902008057, Accu
racy = 0.9247087240219116
Iter #6660096:  Learning rate = 0.000338:   Batch Loss = 0.421057, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4697577953338623, Accu
racy = 0.9400104284286499
Iter #6664192:  Learning rate = 0.000338:   Batch Loss = 0.453292, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.47219154238700867, Acc
uracy = 0.9387932419776917
Iter #6668288:  Learning rate = 0.000338:   Batch Loss = 0.450596, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.4865446984767914, Accu
racy = 0.9316640496253967
Iter #6672384:  Learning rate = 0.000338:   Batch Loss = 0.433707, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.48268651962280273, Acc
uracy = 0.9290558099746704
Iter #6676480:  Learning rate = 0.000338:   Batch Loss = 0.395992, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.48073914647102356, Acc
uracy = 0.9334028959274292
Iter #6680576:  Learning rate = 0.000338:   Batch Loss = 0.442772, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.4723953604698181, Accu
```

```
racy = 0.9361850023269653
Iter #6684672:  Learning rate = 0.000338:   Batch Loss = 0.456496, Accurac
y = 0.951171875
PERFORMANCE ON TEST SET:             Batch Loss = 0.47172361612319946, Acc
uracy = 0.9354894757270813
Iter #6688768:  Learning rate = 0.000338:   Batch Loss = 0.432632, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.4893532395362854, Accu
racy = 0.9297513365745544
Iter #6692864:  Learning rate = 0.000338:   Batch Loss = 0.425106, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.47167110443115234, Acc
uracy = 0.9353156089782715
Iter #6696960:  Learning rate = 0.000338:   Batch Loss = 0.437546, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4762731194496155, Accu
racy = 0.9318379163742065
Iter #6701056:  Learning rate = 0.000324:   Batch Loss = 0.402216, Accurac
y = 0.96484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.47363904118537903, Acc
uracy = 0.9393149018287659
Iter #6705152:  Learning rate = 0.000324:   Batch Loss = 0.444746, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.46923837065696716, Acc
uracy = 0.9363588690757751
Iter #6709248:  Learning rate = 0.000324:   Batch Loss = 0.422576, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.47649168968200684, Acc
uracy = 0.9349678158760071
Iter #6713344:  Learning rate = 0.000324:   Batch Loss = 0.441049, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.48506689071655273, Acc
uracy = 0.9353156089782715
Iter #6717440:  Learning rate = 0.000324:   Batch Loss = 0.375958, Accurac
y = 0.978515625
PERFORMANCE ON TEST SET:             Batch Loss = 0.47026699781417847, Acc
uracy = 0.9367066621780396
Iter #6721536:  Learning rate = 0.000324:   Batch Loss = 0.435317, Accurac
y = 0.955078125
PERFORMANCE ON TEST SET:             Batch Loss = 0.47082674503326416, Acc
uracy = 0.9339245557785034
Iter #6725632:  Learning rate = 0.000324:   Batch Loss = 0.442027, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.47531524300575256, Acc
uracy = 0.9361850023269653
Iter #6729728:  Learning rate = 0.000324:   Batch Loss = 0.424630, Accurac
y = 0.94921875
PERFORMANCE ON TEST SET:             Batch Loss = 0.47933024168014526, Acc
uracy = 0.9332290291786194
Iter #6733824:  Learning rate = 0.000324:   Batch Loss = 0.436801, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.4795084595680237, Accu
racy = 0.9365327954292297
Iter #6737920:  Learning rate = 0.000324:   Batch Loss = 0.441604, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.47667160630226135, Acc
uracy = 0.9351417422294617
Iter #6742016:  Learning rate = 0.000324:   Batch Loss = 0.455511, Accurac
y = 0.943359375
PERFORMANCE ON TEST SET:             Batch Loss = 0.4873870015144348, Accu
racy = 0.9297513365745544
```

```
Iter #6746112:  Learning rate = 0.000324:   Batch Loss = 0.431150, Accurac
y = 0.96484375
PERFORMANCE ON TEST SET:             Batch Loss = 0.484703928232193, Accur
acy = 0.9302729964256287
Iter #6750208:  Learning rate = 0.000324:   Batch Loss = 0.421037, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.4833904206752777, Accu
racy = 0.9327073693275452
Iter #6754304:  Learning rate = 0.000324:   Batch Loss = 0.430378, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.47095876932144165, Acc
uracy = 0.9363588690757751
Iter #6758400:  Learning rate = 0.000324:   Batch Loss = 0.420464, Accurac
y = 0.95703125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4710622727870941, Accu
racy = 0.9387932419776917
Iter #6762496:  Learning rate = 0.000324:   Batch Loss = 0.439891, Accurac
y = 0.947265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.48006874322891235, Acc
uracy = 0.934272289276123
Iter #6766592:  Learning rate = 0.000324:   Batch Loss = 0.414409, Accurac
y = 0.97265625
PERFORMANCE ON TEST SET:             Batch Loss = 0.47117555141448975, Acc
uracy = 0.9344461560249329
Iter #6770688:  Learning rate = 0.000324:   Batch Loss = 0.425943, Accurac
y = 0.9609375
PERFORMANCE ON TEST SET:             Batch Loss = 0.47087913751602173, Acc
uracy = 0.934272289276123
Iter #6774784:  Learning rate = 0.000324:   Batch Loss = 0.434474, Accurac
y = 0.953125
PERFORMANCE ON TEST SET:             Batch Loss = 0.46955859661102295, Acc
uracy = 0.9386193752288818
Iter #6778880:  Learning rate = 0.000324:   Batch Loss = 0.445946, Accurac
y = 0.9453125
PERFORMANCE ON TEST SET:             Batch Loss = 0.4709849953651428, Accu
racy = 0.932881236076355
Iter #6782976:  Learning rate = 0.000324:   Batch Loss = 0.419295, Accurac
y = 0.958984375
PERFORMANCE ON TEST SET:             Batch Loss = 0.47014737129211426, Acc
uracy = 0.9398365616798401
Iter #6787072:  Learning rate = 0.000324:   Batch Loss = 0.439174, Accurac
y = 0.935546875
PERFORMANCE ON TEST SET:             Batch Loss = 0.4663558900356293, Accu
racy = 0.938445508480072
Optimization Finished!
FINAL RESULT: Batch Loss = 0.4663558900356293, Accuracy = 0.93844550848007
2
TOTAL TIME:  7346.106528043747
```

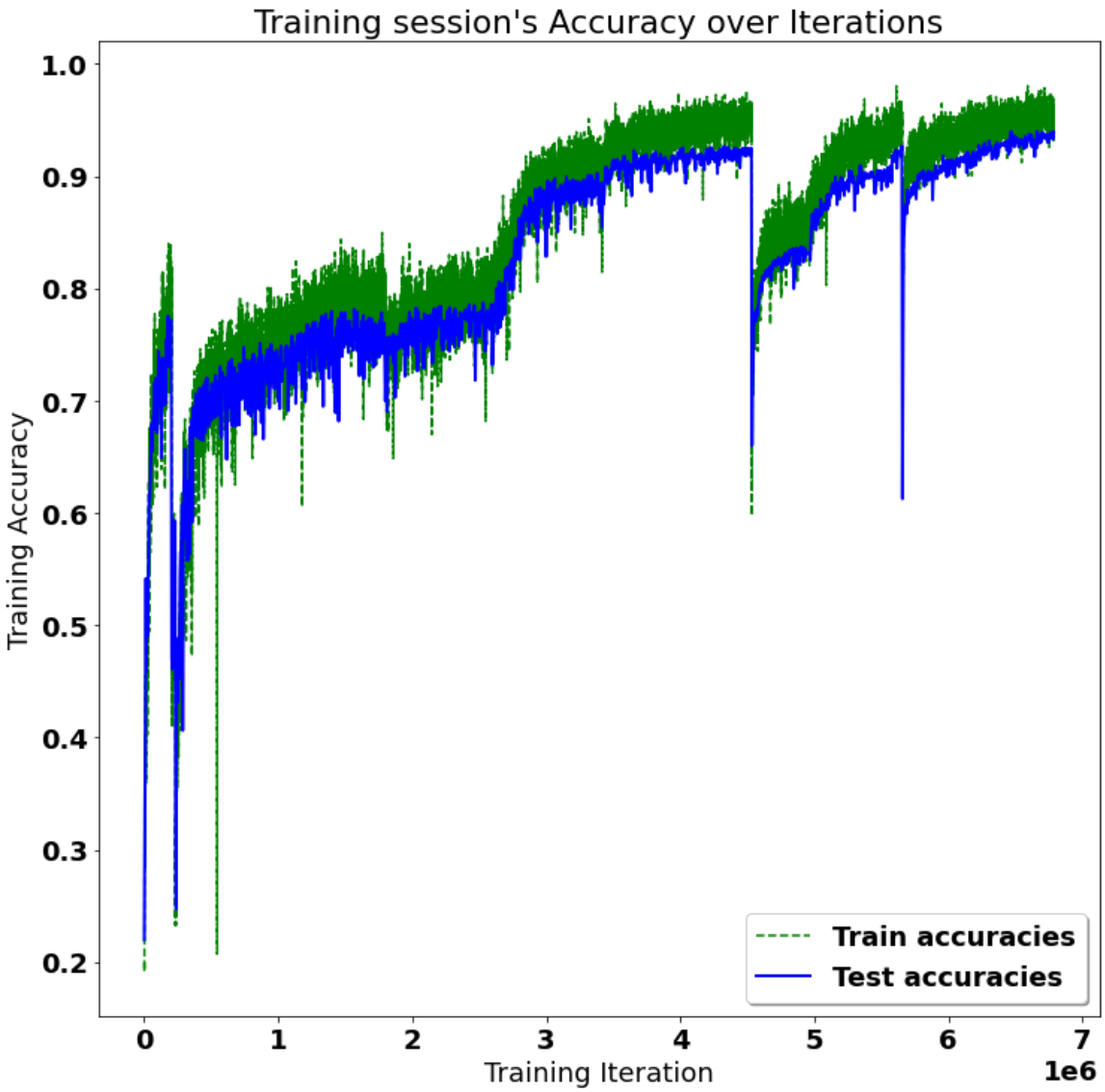
Results:

In [12]:

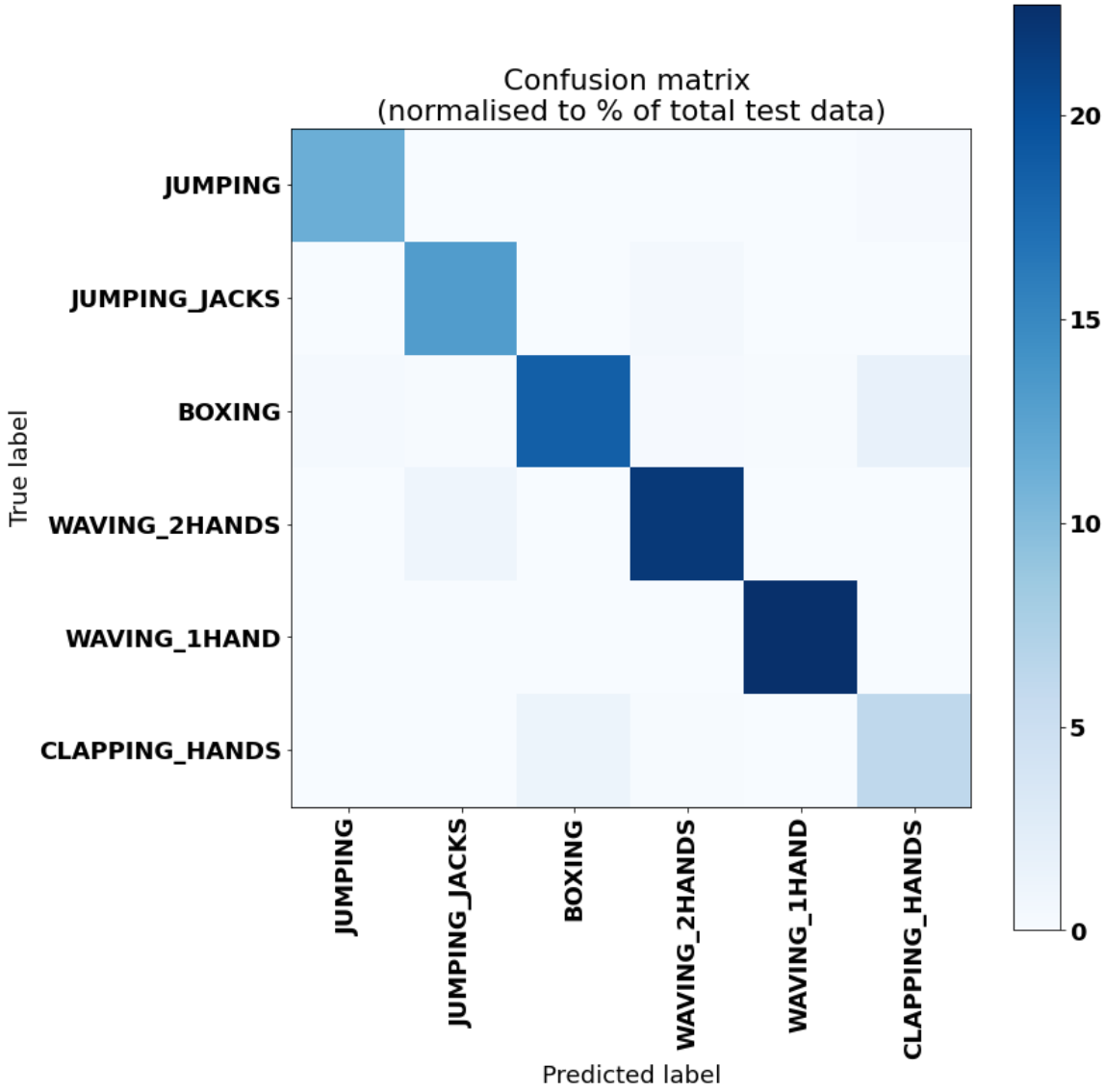
```
# (Inline plots: )
%matplotlib inline

font = {
    'family' : 'DejaVu Sans',
    'weight' : 'bold',
    'size'   : 18
}
matplotlib.rc('font', **font)

width = 12
height = 12
plt.figure(figsize=(width, height))

indep_train_axis = np.array(range(batch_size, (len(train_losses)+1)*batch_size, batch_s
ize))
#plt.plot(indep_train_axis, np.array(train_losses),     "b--", label="Train losses")
plt.plot(indep_train_axis, np.array(train_accuracies), "g--", label="Train accuracies")

indep_test_axis = np.append(
    np.array(range(batch_size, len(test_losses)*display_iter, display_iter)[:-1]),
    [training_iters]
)
#plt.plot(indep_test_axis, np.array(test_losses), "b-", linewidth=2.0, label="Test loss
es")
plt.plot(indep_test_axis, np.array(test_accuracies), "b-", linewidth=2.0, label="Test a
ccuracies")
print(len(test_accuracies))
print(len(train_accuracies))

plt.title("Training session's Accuracy over Iterations")
plt.legend(loc='lower right', shadow=True)
plt.ylabel('Training Accuracy')
plt.xlabel('Training Iteration')

plt.show()

# Results

predictions = one_hot_predictions.argmax(1)

print("Testing Accuracy: {}%".format(100*accuracy))

print("")
print("Precision: {}%".format(100*metrics.precision_score(y_test, predictions, average=
"weighted")))
print("Recall: {}%".format(100*metrics.recall_score(y_test, predictions, average="weigh
ted")))
print("f1_score: {}%".format(100*metrics.f1_score(y_test, predictions, average="weighte
d")))

print("")
print("Confusion Matrix:")
print("Created using test set of {} datapoints, normalised to % of each class in the te
st dataset".format(len(y_test)))
confusion_matrix = metrics.confusion_matrix(y_test, predictions)


#print(confusion_matrix)
```

```
normalised_confusion_matrix = np.array(confusion_matrix, dtype=np.float32)/np.sum(confu
sion_matrix)*100


# Plot Results:
width = 12
height = 12
plt.figure(figsize=(width, height))
plt.imshow(
    normalised_confusion_matrix,
    interpolation='nearest',
    cmap=plt.cm.Blues
)
plt.title("Confusion matrix \n(normalised to % of total test data)")
plt.colorbar()
tick_marks = np.arange(n_classes)
plt.xticks(tick_marks, LABELS, rotation=90)
plt.yticks(tick_marks, LABELS)
plt.tight_layout()
plt.ylabel('True label')
plt.xlabel('Predicted label')
plt.show()
```

```
1659
13256
```




```
Testing Accuracy: 93.8445508480072%

Precision: 93.92252364700282%
Recall: 93.84454877412624%
f1_score: 93.8566063331221%

Confusion Matrix:
Created using test set of 5751 datapoints, normalised to % of each class i
n the test dataset
```



In [13]:

```
X_val_path = DATASET_PATH + "X_val.txt"
X_val = load_X(X_val_path)
print(X_val)

preds = sess.run(
    [pred],
    feed_dict={
        x: X_val
   }
)

print(preds)
```

```
[[[307.589 162.976 319.364 ...   0.    329.752 161.651]
  [307.567 162.979 319.362 ...   0.    328.527 161.655]
  [306.298 162.951 319.351 ...   0.    328.495 161.681]
  ...
  [293.291 122.534 307.676 ... 128.953 315.438 119.884]
  [289.392 140.743 307.615 ...   0.    315.393 139.408]
  [295.848 161.658 307.628 ... 160.331 314.112 160.264]]]
[array([[ 5.2737584 , -1.9157512 ,  0.22798854, -2.9608133 , -1.660554  ,
         0.86934733]], dtype=float32)]
```

In [15]:

```
#sess.close()
print(test_accuracies)
```

[0.21978787, 0.34307078, 0.4336637, 0.54164493, 0.45852897, 0.5367762, 0.4
905234, 0.5263432, 0.5722483, 0.6167623, 0.6131108, 0.62980354, 0.6565814
6, 0.6784907, 0.6548426, 0.6701443, 0.68579376, 0.6877065, 0.69918275, 0.7
160494, 0.6746653, 0.72004867, 0.71013737, 0.6715354, 0.689967, 0.7089202,
0.74404454, 0.7292645, 0.71257174, 0.7348287, 0.7040515, 0.6491045, 0.7304
817, 0.7381325, 0.6974439, 0.7115284, 0.69605285, 0.7221353, 0.6967484, 0.
7209181, 0.7541297, 0.73500264, 0.76073724, 0.7753434, 0.753782, 0.757607
4, 0.72996, 0.76577985, 0.7703008, 0.771518, 0.47087464, 0.46078944, 0.471
74403, 0.5433838, 0.59294033, 0.55972874, 0.59276646, 0.32116154, 0.247261
35, 0.44983482, 0.4682664, 0.43157712, 0.48495913, 0.48913231, 0.45244306,
0.5425143, 0.56564075, 0.56477135, 0.58181185, 0.6077204, 0.40619022, 0.56
63363, 0.6271953, 0.6318901, 0.6567553, 0.6504956, 0.5772909, 0.6016345,
0.62858635, 0.5574683, 0.59276646, 0.5884194, 0.5633803, 0.632064, 0.67657
8, 0.6270214, 0.68301165, 0.5917232, 0.67170924, 0.6746653, 0.6944879, 0.6
685794, 0.6957051, 0.6922274, 0.7061381, 0.7009216, 0.69170576, 0.6668405
5, 0.68022954, 0.6887498, 0.69953054, 0.7111807, 0.6967484, 0.66527563, 0.
7064858, 0.7005738, 0.70022607, 0.664754, 0.713615, 0.6934446, 0.69118416,
0.6715354, 0.7179621, 0.7120501, 0.72039646, 0.7123979, 0.68335944, 0.6772
735, 0.6837072, 0.69327074, 0.68179446, 0.707703, 0.68266386, 0.71535385,
0.69918275, 0.6977917, 0.6937924, 0.69796556, 0.68492436, 0.7170927, 0.722
6569, 0.69831336, 0.695879, 0.6868371, 0.7080508, 0.7170927, 0.7087463, 0.
70387757, 0.7304817, 0.6704921, 0.6670144, 0.6704921, 0.7249174, 0.7181359
5, 0.7118762, 0.72126585, 0.7191793, 0.71761435, 0.6915319, 0.71048516, 0.
6477134, 0.7205703, 0.7040515, 0.72996, 0.71726656, 0.734481, 0.6936185,
0.727004, 0.7257868, 0.7369153, 0.6776213, 0.72248304, 0.70422536, 0.71222
395, 0.72961223, 0.6896192, 0.67709965, 0.72039646, 0.6937924, 0.7238741,
0.6998783, 0.69831336, 0.6950096, 0.67831683, 0.7330899, 0.72874284, 0.729
61223, 0.70109546, 0.7261346, 0.73239434, 0.7155277, 0.74769604, 0.7358720
3, 0.710659, 0.6930969, 0.7308294, 0.7330899, 0.73569816, 0.7080508, 0.688
402, 0.73030776, 0.6997044, 0.7280473, 0.73552424, 0.6863154, 0.73239434,
0.72874284, 0.69240135, 0.7407407, 0.74161017, 0.73587203, 0.738828, 0.669
7966, 0.7308294, 0.70509475, 0.72439575, 0.7316988, 0.72648233, 0.7374369
5, 0.69240135, 0.7431751, 0.7509998, 0.7332638, 0.6804034, 0.74056685, 0.7
3569816, 0.7089202, 0.66579723, 0.71013737, 0.72352636, 0.73343766, 0.7271
7786, 0.7398713, 0.7290906, 0.7457833, 0.7143105, 0.7400452, 0.72717786,
0.69327074, 0.7099635, 0.7475222, 0.7501304, 0.7421318, 0.71570164, 0.7435
229, 0.73500264, 0.73778474, 0.71048516, 0.740393, 0.7423057, 0.7155277,
0.74491394, 0.7370892, 0.7339593, 0.7320466, 0.69883496, 0.7372631, 0.6927
491, 0.7442184, 0.7489132, 0.7457833, 0.71726656, 0.72352636, 0.7313511,
0.72961223, 0.73256826, 0.7332638, 0.7313511, 0.7473483, 0.7490871, 0.7407
407, 0.74769604, 0.69101024, 0.7395236, 0.7155277, 0.7438706, 0.7363937,
0.7442184, 0.74682665, 0.75778127, 0.7146583, 0.7497826, 0.74682665, 0.727
3518, 0.7186576, 0.7360459, 0.6974439, 0.7383064, 0.75273865, 0.74526167,
0.75899845, 0.75482523, 0.73552424, 0.7520431, 0.7509998, 0.7600417, 0.753
08645, 0.71048516, 0.73378545, 0.7520431, 0.7475222, 0.75586855, 0.7595200
5, 0.6990089, 0.7310033, 0.7605634, 0.7541297, 0.76804036, 0.7544775, 0.75
91723, 0.7563902, 0.7226569, 0.7202226, 0.76421493, 0.7381325, 0.7572596,
0.7650843, 0.76386714, 0.76977915, 0.7485655, 0.72596073, 0.7633455, 0.769
77915, 0.7600417, 0.75552076, 0.759694, 0.752217, 0.74056685, 0.76386714,
0.7276995, 0.7690836, 0.73500264, 0.71952707, 0.7753434, 0.7749956, 0.7344
81, 0.69483566, 0.7556947, 0.75865066, 0.7308294, 0.77638674, 0.77708226,
0.7494349, 0.75552076, 0.73865414, 0.76421493, 0.771518, 0.7198748, 0.7696
053, 0.7450878, 0.7706486, 0.7795166, 0.73239434, 0.74143624, 0.7443923,
0.71639717, 0.7410885, 0.7553469, 0.68927145, 0.7628239, 0.7593462, 0.7659
537, 0.7351765, 0.7725613, 0.68179446, 0.76421493, 0.7395236, 0.77395236,
0.77169186, 0.7784733, 0.73934966, 0.7426534, 0.7704747, 0.74769604, 0.754
3036, 0.7781255, 0.77829945, 0.7525648, 0.76577985, 0.75047815, 0.753782,
0.75047815, 0.74282736, 0.77117026, 0.76264995, 0.7569118, 0.7541297, 0.74
94349, 0.764041, 0.74995655, 0.74491394, 0.76108503, 0.7814293, 0.773083,
0.7654321, 0.7737785, 0.774474, 0.74456614, 0.77916884, 0.74804384, 0.7543
036, 0.7569118, 0.7221353, 0.7636933, 0.75795513, 0.76143277, 0.77099633,

```
0.77082247, 0.74334896, 0.7758651, 0.7191793, 0.7703008, 0.7471744, 0.7668
232, 0.73587203, 0.76299775, 0.7489132, 0.7706486, 0.7675187, 0.7736046,
0.73430705, 0.7501304, 0.76804036, 0.7593462, 0.77290905, 0.7664754, 0.754
65137, 0.72961223, 0.74804384, 0.74960876, 0.7313511, 0.76890975, 0.738480
27, 0.76212835, 0.7473483, 0.76143277, 0.7663015, 0.74647886, 0.73865414,
0.7713441, 0.75186926, 0.74804384, 0.7781255, 0.7419579, 0.75239086, 0.758
99845, 0.76456267, 0.7690836, 0.7391758, 0.6998783, 0.7188315, 0.75586855,
0.6903147, 0.7186576, 0.7489132, 0.7442184, 0.75708574, 0.7562163, 0.74195
79, 0.7461311, 0.749261, 0.7410885, 0.7544775, 0.7593462, 0.7330899, 0.739
69746, 0.7525648, 0.7028343, 0.7576074, 0.75360805, 0.7543036, 0.75465137,
0.74560946, 0.7543036, 0.75465137, 0.759694, 0.76386714, 0.7683881, 0.7476
9604, 0.7612589, 0.76021564, 0.7643888, 0.7683881, 0.7623022, 0.73343766,
0.72161365, 0.759694, 0.7633455, 0.7690836, 0.74995655, 0.75360805, 0.7400
452, 0.7736046, 0.77882105, 0.77516955, 0.75082594, 0.7501304, 0.77325684,
0.7572596, 0.7485655, 0.7565641, 0.7593462, 0.7652582, 0.75482523, 0.74265
34, 0.7826465, 0.7471744, 0.7703008, 0.72648233, 0.7793427, 0.7623022, 0.7
703008, 0.7532603, 0.771518, 0.7687359, 0.76769257, 0.7706486, 0.73500264,
0.7623022, 0.72874284, 0.77743, 0.76577985, 0.7654321, 0.74769604, 0.74839
157, 0.7520431, 0.7758651, 0.7722135, 0.75465137, 0.7723874, 0.7663015, 0.
7482177, 0.77638674, 0.76091117, 0.7461311, 0.7336115, 0.746305, 0.747869
9, 0.78195095, 0.773083, 0.76734483, 0.755173, 0.7668232, 0.7633455, 0.773
95236, 0.774474, 0.783342, 0.7723874, 0.76786643, 0.773083, 0.78403753, 0.
7623022, 0.74839157, 0.76491046, 0.74491394, 0.7706486, 0.7699531, 0.75465
137, 0.7541297, 0.771518, 0.7605634, 0.7584768, 0.7796905, 0.771518, 0.776
5606, 0.7443923, 0.7829943, 0.78003824, 0.7513476, 0.7628239, 0.76073724,
0.7765606, 0.768562, 0.7749956, 0.77169186, 0.74769604, 0.78195095, 0.7584
768, 0.7619544, 0.76299775, 0.77099633, 0.77290905, 0.77082247, 0.7490871,
0.7633455, 0.7784733, 0.774474, 0.77117026, 0.7544775, 0.77290905, 0.77325
684, 0.7765606, 0.7777778, 0.77290905, 0.77395236, 0.7812554, 0.7796905,
0.77203965, 0.7814293, 0.77082247, 0.7798644, 0.78108156, 0.78542864, 0.75
708574, 0.7758651, 0.76143277, 0.7786472, 0.7777778, 0.7737785, 0.7796905,
0.77203965, 0.784907, 0.7758651, 0.71813595, 0.7633455, 0.76612765, 0.7730
83, 0.76091117, 0.77829945, 0.7683881, 0.7574335, 0.7569118, 0.77882105,
0.7696053, 0.7798644, 0.7718658, 0.77829945, 0.78716743, 0.78351593, 0.752
73865, 0.7619544, 0.7438706, 0.7543036, 0.74995655, 0.7793427, 0.78038603,
0.7798644, 0.77117026, 0.7795166, 0.76456267, 0.7633455, 0.78108156, 0.785
6025, 0.7845592, 0.73187274, 0.7442184, 0.7746479, 0.7737785, 0.784907, 0.
7809077, 0.78786296, 0.7699531, 0.8002087, 0.759694, 0.7845592, 0.7864719,
0.80594677, 0.79307944, 0.7650843, 0.7998609, 0.7889063, 0.7786472, 0.7725
613, 0.8045557, 0.7974265, 0.8017736, 0.77638674, 0.80942446, 0.8167275,
0.81116325, 0.81185883, 0.818988, 0.80907667, 0.8076856, 0.81533647, 0.798
8176, 0.81864023, 0.82281345, 0.8210746, 0.80559903, 0.8266388, 0.8080334,
0.7986437, 0.80872893, 0.8313337, 0.841245, 0.8447226, 0.8388106, 0.825421
7, 0.8292471, 0.84281, 0.8365502, 0.8309859, 0.8624587, 0.8450704, 0.82559
556, 0.83985394, 0.8391584, 0.84176666, 0.8657625, 0.847157, 0.8676752, 0.
86976177, 0.87410885, 0.8603721, 0.8706312, 0.87132674, 0.86384976, 0.8648
931, 0.86471915, 0.84854805, 0.8513302, 0.8810642, 0.8699357, 0.87132674,
0.86193705, 0.87619543, 0.8589811, 0.8695879, 0.86071986, 0.868023, 0.8702
834, 0.85515565, 0.888715, 0.8754999, 0.8615893, 0.841245, 0.8577639, 0.88
36724, 0.86611027, 0.8845418, 0.8901061, 0.87532604, 0.8848896, 0.8855851,
0.8841941, 0.87219614, 0.87602156, 0.8873239, 0.8572422, 0.8735872, 0.8786
298, 0.82837766, 0.87689096, 0.8786298, 0.8958442, 0.8542862, 0.8838463,
0.88402015, 0.8603721, 0.89149714, 0.89306206, 0.8765432, 0.8786298, 0.894
8009, 0.8873239, 0.88628066, 0.8970614, 0.8979308, 0.8706312, 0.8688924,
0.87132674, 0.87410885, 0.85046077, 0.8880195, 0.88749784, 0.8880195, 0.88
123804, 0.8688924, 0.89219266, 0.8899322, 0.8788037, 0.89027995, 0.887323
9, 0.8748044, 0.8904538, 0.88715005, 0.8939315, 0.88228136, 0.89932185, 0.
87915146, 0.8873239, 0.8928882, 0.8779343, 0.89532256, 0.8765432, 0.881933
57, 0.85289514, 0.86628413, 0.8838463, 0.88506347, 0.8968875, 0.89149714,
0.8767171, 0.8954964, 0.8775865, 0.88315076, 0.8920188, 0.8868023, 0.89514
87, 0.8786298, 0.8796731, 0.89462703, 0.8895844, 0.87532604, 0.89306206,
```

```
0.893236, 0.8814119, 0.88123804, 0.8796731, 0.8967136, 0.89653975, 0.88367
24, 0.87323946, 0.89184487, 0.8751522, 0.88210744, 0.88836724, 0.8998435,
0.9008868, 0.88819337, 0.88193357, 0.89254045, 0.8833246, 0.9007129, 0.874
2827, 0.89740914, 0.89619195, 0.8935837, 0.88819337, 0.8817597, 0.885759,
0.8948009, 0.88349855, 0.88506347, 0.9026256, 0.88941056, 0.89254045, 0.87
42827, 0.888715, 0.86784905, 0.8869762, 0.87254393, 0.8941054, 0.854634,
0.8770649, 0.8748044, 0.87723875, 0.8829769, 0.8937576, 0.9010607, 0.90836
376, 0.9010607, 0.9040167, 0.8970614, 0.89306206, 0.8951487, 0.90019125,
0.9057555, 0.8967136, 0.9022779, 0.9017562, 0.9113198, 0.9118414, 0.909059
3, 0.91462356, 0.90558165, 0.910972, 0.9099287, 0.9062772, 0.90749437, 0.9
151452, 0.91149366, 0.90923315, 0.9099287, 0.8963658, 0.9069727, 0.912537,
0.9116675, 0.91149366, 0.90645105, 0.91271085, 0.9101026, 0.91671014, 0.91
81012, 0.9099287, 0.8988002, 0.8941054, 0.90053904, 0.9066249, 0.88941056,
0.9134064, 0.90836376, 0.90332115, 0.91045034, 0.9088854, 0.9099287, 0.894
8009, 0.8824552, 0.90366894, 0.90958095, 0.90836376, 0.92262214, 0.912189
2, 0.9094071, 0.91984004, 0.9113198, 0.9066249, 0.9121892, 0.90645105, 0.9
085376, 0.9116675, 0.9113198, 0.9134064, 0.916884, 0.908016, 0.89427924,
0.9062772, 0.90836376, 0.918449, 0.91149366, 0.9062772, 0.9156668, 0.91427
577, 0.91949224, 0.91549295, 0.9149713, 0.8986263, 0.89462703, 0.89445317,
0.9102765, 0.9088854, 0.91462356, 0.88628066, 0.9113198, 0.91862285, 0.913
7541, 0.9191445, 0.91862285, 0.9121892, 0.9141019, 0.9215789, 0.91271085,
0.910972, 0.9027995, 0.9147974, 0.90679884, 0.91062427, 0.9160146, 0.90523
386, 0.9196662, 0.9172318, 0.9128847, 0.9121892, 0.91462356, 0.91149366,
0.9149713, 0.91584074, 0.91271085, 0.91671014, 0.9137541, 0.910972, 0.9048
861, 0.9066249, 0.9132325, 0.90679884, 0.9022779, 0.90366894, 0.9031473,
0.91045034, 0.9113198, 0.9181012, 0.9161885, 0.91757953, 0.9165363, 0.9167
1014, 0.91444963, 0.91114587, 0.90749437, 0.90436447, 0.91271085, 0.915840
74, 0.92140496, 0.91862285, 0.9212311, 0.91584074, 0.916884, 0.9241871, 0.
9179273, 0.9219266, 0.9149713, 0.9252304, 0.9141019, 0.90749437, 0.903495
1, 0.9187967, 0.8937576, 0.91584074, 0.9193184, 0.91636235, 0.91062427, 0.
91949224, 0.9121892, 0.9118414, 0.9149713, 0.9172318, 0.92175275, 0.920187
8, 0.9221005, 0.92453486, 0.91705793, 0.9219266, 0.9141019, 0.91427577, 0.
9057555, 0.9179273, 0.9156668, 0.9160146, 0.9156668, 0.91775346, 0.919318
4, 0.918449, 0.91827506, 0.91045034, 0.9196662, 0.9101026, 0.9149713, 0.91
236305, 0.9179273, 0.9085376, 0.90558165, 0.9156668, 0.91462356, 0.9174056
6, 0.92383933, 0.91775346, 0.92262214, 0.9191445, 0.91079813, 0.9134064,
0.91271085, 0.9224483, 0.92575204, 0.9193184, 0.9181012, 0.91549295, 0.918
7967, 0.9224483, 0.91462356, 0.9128847, 0.9137541, 0.9200139, 0.9172318,
0.9165363, 0.9022779, 0.92175275, 0.91949224, 0.9099287, 0.91584074, 0.924
1871, 0.924361, 0.91584074, 0.9196662, 0.9222744, 0.9231438, 0.92366546,
0.92662144, 0.9219266, 0.91271085, 0.91705793, 0.9231438, 0.91671014, 0.92
12311, 0.9259259, 0.9193184, 0.9203617, 0.91444963, 0.9156668, 0.92296994,
0.91984004, 0.91949224, 0.91358024, 0.91444963, 0.9151452, 0.9147974, 0.92
453486, 0.9200139, 0.92488265, 0.9147974, 0.9196662, 0.91775346, 0.9127108
5, 0.92349154, 0.91862285, 0.92070943, 0.90732044, 0.91584074, 0.9233177,
0.92349154, 0.9210572, 0.92488265, 0.924361, 0.91862285, 0.92488265, 0.923
83933, 0.922796, 0.91775346, 0.92488265, 0.9215789, 0.92366546, 0.9245348
6, 0.91984004, 0.918449, 0.9241871, 0.9233177, 0.92262214, 0.92488265, 0.6
600591, 0.72874284, 0.7254391, 0.7643888, 0.77412623, 0.7734307, 0.7788210
5, 0.771518, 0.77099633, 0.7816032, 0.7758651, 0.7836898, 0.78786296, 0.79
65571, 0.7955138, 0.8036863, 0.80351245, 0.80542517, 0.7976004, 0.8043818
5, 0.8113372, 0.8073379, 0.8073379, 0.8095983, 0.81220657, 0.81551033, 0.8
132499, 0.81220657, 0.8170753, 0.8149887, 0.81116325, 0.8132499, 0.8177708
4, 0.8198574, 0.8198574, 0.8073379, 0.8191619, 0.8163798, 0.8226395, 0.824
0306, 0.82211787, 0.82020515, 0.8250739, 0.82281345, 0.82020515, 0.824030
6, 0.82455224, 0.821944, 0.8275083, 0.8226395, 0.8254217, 0.82681274, 0.82
26395, 0.8210746, 0.816032, 0.81933576, 0.8250739, 0.82802993, 0.821944,
0.8231612, 0.82872546, 0.83168143, 0.8226395, 0.8212485, 0.82455224, 0.823
68284, 0.8226395, 0.8243784, 0.83394194, 0.82994264, 0.82994264, 0.834115
8, 0.8313337, 0.8297687, 0.835333, 0.8292471, 0.7998609, 0.8129021, 0.8090
7667, 0.82646495, 0.83585465, 0.821944, 0.8262911, 0.83550686, 0.8332464,
```

```
0.8222918, 0.8328986, 0.82159626, 0.8313337, 0.8365502, 0.83237696, 0.8370
7184, 0.83359414, 0.8328986, 0.823509, 0.82768214, 0.8362024, 0.82837766,
0.83150756, 0.83046424, 0.83185536, 0.8320292, 0.8313337, 0.83168143, 0.83
081204, 0.8292471, 0.82524776, 0.8384629, 0.85289514, 0.85689443, 0.866284
13, 0.8586333, 0.8534168, 0.8615893, 0.86419755, 0.8648931, 0.8704573, 0.8
5202575, 0.8692401, 0.86819685, 0.8574161, 0.8695879, 0.86611027, 0.864197
55, 0.86663187, 0.8794992, 0.8683707, 0.86541474, 0.87219614, 0.8685446,
0.86941403, 0.8754999, 0.87289166, 0.8735872, 0.87219614, 0.8702834, 0.876
3693, 0.88315076, 0.8704573, 0.8767171, 0.88402015, 0.8735872, 0.88315076,
0.88315076, 0.8836724, 0.8909755, 0.8920188, 0.8935837, 0.8888889, 0.88767
17, 0.88436794, 0.8869762, 0.89932185, 0.8880195, 0.8901061, 0.8746305, 0.
89114934, 0.8941054, 0.88906276, 0.8935837, 0.88123804, 0.8888889, 0.89880
02, 0.8873239, 0.89845246, 0.8937576, 0.88819337, 0.8937576, 0.8826291, 0.
8939315, 0.8941054, 0.89062774, 0.8892366, 0.89340985, 0.8967136, 0.894453
17, 0.8864545, 0.8909755, 0.9048861, 0.8963658, 0.8960181, 0.9040167, 0.89
48009, 0.90053904, 0.89445317, 0.89619195, 0.89340985, 0.86941403, 0.88958
44, 0.893236, 0.8868023, 0.8939315, 0.8847157, 0.89966965, 0.8923665, 0.89
79308, 0.89462703, 0.90958095, 0.8988002, 0.89532256, 0.90366894, 0.892888
2, 0.8899322, 0.89219266, 0.9031473, 0.8970614, 0.89740914, 0.9017562, 0.9
029734, 0.90019125, 0.8986263, 0.89845246, 0.9008868, 0.8986263, 0.8982785
3, 0.8989741, 0.90332115, 0.8994957, 0.9040167, 0.9031473, 0.90123457, 0.9
000174, 0.8968875, 0.89254045, 0.88436794, 0.8951487, 0.89114934, 0.897409
14, 0.90523386, 0.8979308, 0.9010607, 0.8967136, 0.9048861, 0.9027995, 0.9
047122, 0.9031473, 0.8876717, 0.89810467, 0.8951487, 0.8979308, 0.9036689
4, 0.9059294, 0.9003652, 0.90140843, 0.90366894, 0.9048861, 0.8904538, 0.9
071466, 0.908016, 0.9057555, 0.9034951, 0.90053904, 0.9066249, 0.90923315,
0.89445317, 0.91271085, 0.9189706, 0.91584074, 0.91271085, 0.918449, 0.918
449, 0.9165363, 0.92349154, 0.91705793, 0.9069727, 0.91462356, 0.9196662,
0.92140496, 0.9212311, 0.9224483, 0.9201878, 0.92296994, 0.91740566, 0.926
96923, 0.6125891, 0.7170927, 0.8341158, 0.8509824, 0.86106765, 0.87254393,
0.8748044, 0.8751522, 0.8668058, 0.8876717, 0.8751522, 0.88280296, 0.87967
31, 0.88089025, 0.8817597, 0.88402015, 0.8897583, 0.8899322, 0.8829769, 0.
89027995, 0.8904538, 0.8935837, 0.89149714, 0.89427924, 0.8885411, 0.87967
31, 0.8928882, 0.8895844, 0.8948009, 0.90123457, 0.89653975, 0.89827853,
0.9017562, 0.897583, 0.9007129, 0.9031473, 0.90140843, 0.9026256, 0.904886
1, 0.9010607, 0.8909755, 0.89775693, 0.8994957, 0.89966965, 0.90210396, 0.
89740914, 0.90158236, 0.90453833, 0.89619195, 0.9040167, 0.90332115, 0.903
32115, 0.89932185, 0.8998435, 0.9027995, 0.8788037, 0.89445317, 0.899148,
0.89810467, 0.90053904, 0.9094071, 0.90019125, 0.9022779, 0.89775693, 0.90
506, 0.9019301, 0.9061033, 0.9040167, 0.90332115, 0.9090593, 0.9029734, 0.
90749437, 0.9078421, 0.9057555, 0.91827506, 0.90836376, 0.9048861, 0.90558
165, 0.9081899, 0.9061033, 0.9141019, 0.89932185, 0.9120153, 0.9041906, 0.
91444963, 0.9008868, 0.9116675, 0.908016, 0.91062427, 0.9121892, 0.9144496
3, 0.9088854, 0.90158236, 0.9034951, 0.9078421, 0.9027995, 0.91358024, 0.9
099287, 0.91427577, 0.9102765, 0.91584074, 0.91114587, 0.9102765, 0.908189
9, 0.9120153, 0.9121892, 0.91705793, 0.91584074, 0.91427577, 0.9156668, 0.
9128847, 0.9141019, 0.9130586, 0.9179273, 0.91705793, 0.91392803, 0.919840
04, 0.9031473, 0.9210572, 0.9215789, 0.91757953, 0.9179273, 0.9187967, 0.9
1149366, 0.90732044, 0.9147974, 0.912537, 0.9118414, 0.91114587, 0.9217527
5, 0.9088854, 0.91427577, 0.91045034, 0.9132325, 0.9099287, 0.91984004, 0.
9203617, 0.91984004, 0.92262214, 0.9274909, 0.91584074, 0.9200139, 0.91444
963, 0.91984004, 0.9233177, 0.92175275, 0.9241871, 0.9271431, 0.9215789,
0.91862285, 0.9294036, 0.922796, 0.92488265, 0.9193184, 0.92070943, 0.9236
6546, 0.9215789, 0.92453486, 0.9314902, 0.92349154, 0.924361, 0.92662144,
0.92488265, 0.9151452, 0.91584074, 0.9224483, 0.9189706, 0.9221005, 0.9201
878, 0.9274909, 0.92731696, 0.91549295, 0.9299252, 0.9187967, 0.9271431,
0.9203617, 0.9264476, 0.92575204, 0.9280125, 0.92731696, 0.9292297, 0.9280
125, 0.9290558, 0.9292297, 0.9274909, 0.9294036, 0.9300991, 0.9271431, 0.9
2731696, 0.9309685, 0.922796, 0.92383933, 0.9255782, 0.92488265, 0.9194922
4, 0.93357676, 0.93044686, 0.9222744, 0.93983656, 0.92696923, 0.9287081,
0.92957747, 0.93253344, 0.93601114, 0.92609984, 0.9353156, 0.9306208, 0.93
```

```
079466, 0.93253344, 0.9337506, 0.9321857, 0.92731696, 0.9259259, 0.9273169
6, 0.9294036, 0.9281864, 0.9330551, 0.93270737, 0.9300991, 0.9290558, 0.93
11424, 0.92488265, 0.92888194, 0.9309685, 0.9353156, 0.92540425, 0.930968
5, 0.92975134, 0.9342723, 0.9346201, 0.9353156, 0.93757606, 0.9334029, 0.9
36185, 0.92853415, 0.93670666, 0.9262737, 0.93079466, 0.93288124, 0.926447
6, 0.9250565, 0.91740566, 0.92662144, 0.9292297, 0.9247087, 0.9400104, 0.9
3879324, 0.93166405, 0.9290558, 0.9334029, 0.936185, 0.9354895, 0.9297513
4, 0.9353156, 0.9318379, 0.9393149, 0.93635887, 0.9349678, 0.9353156, 0.93
670666, 0.93392456, 0.936185, 0.933229, 0.9365328, 0.93514174, 0.92975134,
0.930273, 0.93270737, 0.93635887, 0.93879324, 0.9342723, 0.93444616, 0.934
2723, 0.9386194, 0.93288124, 0.93983656, 0.9384455, 0.9384455]
```

Conclusion

Final accuracy of >90% is pretty good, considering that training takes about 7 minutes.

Noticeable confusion between activities of Clapping Hands and Boxing, and between Jumping Jacks and Waving Two Hands which is understandable.

In terms of the applicability of this to a wider dataset, I would imagine that it would be able to work for any activities in which the training included a views from all angles to be tested on. It would be interesting to see it's applicability to camera angles in between the 4 used in this dataset, without training on them specifically.

Overall, this experiment validates the idea that 2D pose can be used for at least human activity recognition, and provides verification to continue onto use of 2D pose for behaviour estimation in both people and animals

With regards to Using LSTM-RNNs

- Batch sampling
  - It is neccessary to ensure you are not just sampling classes one at a time! (ie y_train is ordered by class and batch chosen in order)The use of random sampling of batches without replacement from the training data resolves this.
- Architecture
  - Testing has been run using a variety of hidden units per LSTM cell, with results showing that testing accuracy achieves a higher score when using a number of hidden cells approximately equal to that of the input, ie 34. The following figure displays the final accuracy achieved on the testing dataset for a variety of hidden units, all using a batch size of 4096 and 300 epochs (a total of 1657 iterations, with testing performed every 8th iteration).

Future Works

Inclusion of :

- A pipeline for qualitative results
- A validation dataset
- Momentum
- Normalise input data (each point with respect to distribution of itself only)
- Dropout
- Comparison of effect of changing batch size

Further research will be made into the use on more subtle activity classes, such as walking versus running, agitated movement versus calm movement, and perhaps normal versus abnormal behaviour, based on a baseline of normal motion.

References

The dataset can be found at http://tele-immersion.citris-uc.org/berkeley_mhad (http://tele-immersion.citris-uc.org/berkeley_mhad) released under the BSD-2 license

> Copyright (c) 2013, Regents of the University of California All rights reserved.

The network used in this experiment is based on the following, available under the MIT License (https://github.com/guillaume-chevalier/LSTM-Human-Activity-Recognition/blob/master/LICENSE). :

> Guillaume Chevalier, LSTMs for Human Activity Recognition, 2016
> https://github.com/guillaume-chevalier/LSTM-Human-Activity-Recognition
> (https://github.com/guillaume-chevalier/LSTM-Human-Activity-Recognition)

In [ ]:

```
# Let's convert this notebook to a README for the GitHub project's title page:
!jupyter nbconvert --to markdown LSTM.ipynb
!mv LSTM.md README.md
```

In [ ]:

```

```