Homework 4 CNN network for the EMNIST data

Modify the code in `Example4_MNIST_CNN.ipynb` to run on EMNIST data file.
You will need to
(1) Load `emnist-balanced.pkl gz` instead of `mnist.pkl.gz`
(2) Change the number of output nodes from 10 to 47.

There are 52 upper and lower case letters in the English Alphabet along with 10 digits for a total of 62 characters.
The data set you are given has only 47 classes because the upper and lower cases for the 15 letters C, I, J, K, L, M, O, P, S, U, V, W, X, Y, Z were merged together.

You should be able to *easily* get over 85% accuracy. Can you get over 90%? Try!

Handin your `.ipynb` flile and a short report 1-2 pages on your results.

Short Report: Report with the following information about your network:
number of layers with the number of neurons for each layer.
Activation functions used on each layer.
Weight initializer used for each layer.
learning rate value (if used).
Mini-batch size used.
Optimizer used.
Loss function used.
dropout value if used.
lmda value for the regularizer if used.
Give the final **test** accuracy.

Explain what you did to get a higher accuracy. Your starting point are the hyperparameter values in `Example4_MNIST_CNN.ipynb` covered in class and posted on blackboard. Keep it short and precise. Specifically,

- Don't use vague/imprecise words in your report. For example, saying the results are "satisfactory" doesn't mean anything as you have not explained what you mean by "satisfactory". Similar problem with using descriptors such as "excellent", "good", "adequate", etc.

- Don't make unsubstantiated claims like " with a proper understanding ... a higher accuracy ... can be achieved". What does "proper understanding" even mean? How can you claim you will get a "higher accuracy"?

- When you say something was "optimized" it means you had some criteria and you found the minimum using that criteria. Don't say you found the "optimum" value unless you state the criteria and then actually found the minimum according to that criteria.