File structure:
- HW1P2
  - main.py
  - 11-785hw2p2-s20
    - provided file and folders including data from kaggle
  - checkpoint
    - ContContInitWeight_BaselineSGD_StepLR_Epoch4.txt
  - output
    - Saved label test file for classification and verification

To run my model:
- Have the file structure as above
- Type "source activate pytorch_p36" in the terminal
- In the HW2P2 directory, type "python3 main.py" to run the model
- After the model finishes, find predicted test_class_labels.npy and csv files under output folder

Network design:
- Used MobileNetV2 to implement, the configuration is described on the right
- Used SGD as optimizer with lr = 0.01 with Nesterov momentum = 0.9 and weight decay = 5e_4 with a batch size = 256
- Used ReduceLROnPlateau scheduler with mode = “min”, factor = 0.5, patience=1
- Used CrossEntropyLoss
- Trained with the above configuration for 55 epochs, then changed the scheduler’s factor=0.8 and trained for 10 more epochs.

Operator	t	c	n	s
conv2d 3x3	-	32	1	1
bottleneck	1	16	1	1
bottleneck	6	24	2	2
bottleneck	6	32	3	1
bottleneck	6	64	4	2
bottleneck	6	96	3	1
bottleneck	6	160	3	1
bottleneck	6	320	1	1
conv2d 1x1				1
avgpool				
flatten				
linear layer				

Reference Links:
- http://www.cs.cmu.edu/~bhiksha/courses/deeplearning/Fall.2019/www.f19/document/note/hwnotes/hw2/hw2.html
- https://arxiv.org/pdf/1801.04381.pdf
- http://deeplearning.cs.cmu.edu/document/recitation/Recitation%206.pdf
- https://github.com/QuantScientist/Deep-Learning-Boot-Camp