1. I. Introduction:

這次LAB的目的主要在於熟悉Data Augmentation 以及Data Sampler 相關API的使用，學會在面對Data imbalanced或是資料過少的問題時，如何有效利用這些方法去提升模型的準確率。

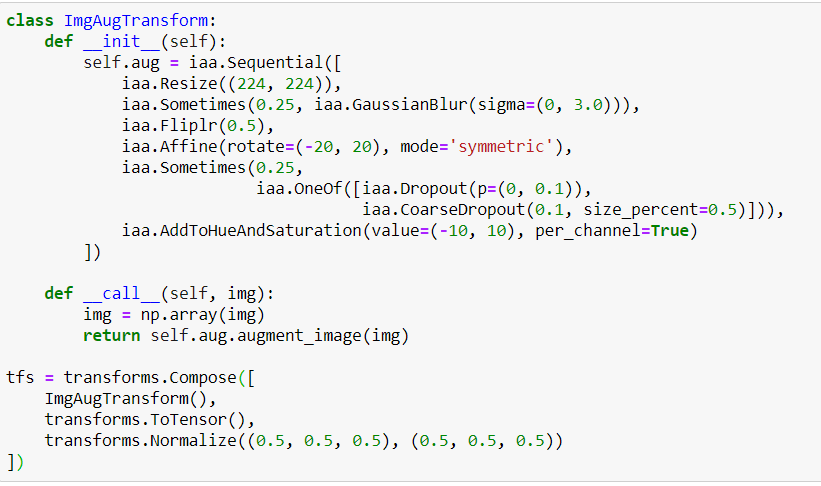
II. Experiment setup

同Lab1

1. III. Result

2-1-1:

以下是我結合imgaug以及torch.transform的data Augmentation 方式



其中16筆資料轉換後的結果

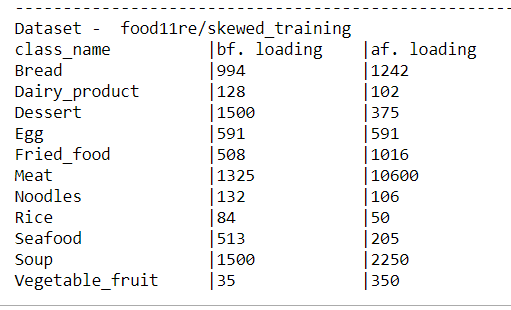


2-1-2

我以100為基準(表示不變)，超過100根據相對於100的倍率增加image\_filenames裡的資料(隨機採樣並複製)，低於100則刪除，使用的權重是原本的例子

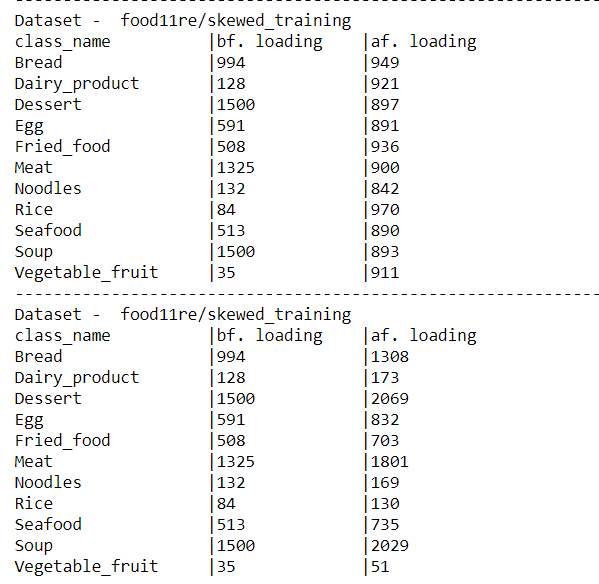
wts = [ 125, 80, 25, 100, 200, 800, 80, 60, 40, 150, 1000 ]

最後結果:



2-2

以每個class的sample數量的倒數作為那個class的權重，最後就能得到平均分配的結果，我採樣10000筆資料，上半為WeightedRandomSampler，下半為RandomSampler，可以發現WeightedRandomSampler後的結果是平均分配的，而RandomSampler的比例還是原本每個class當中sample數量的比例



2-3

來不及train完=ˇ=(LAB1用resnet18當pretrained model有到86%)