W3H1 Image Classification

Given an image, we aim to build a neural network to predict the object. For example, the following image has the label 'cat'.

Data:

Download image dataset and sample code:

- Training set: 9867 jpg images
- Validation set: 3431 jpg images
- Test set: 1501 jpg images
- Label: 11 major food categories including bread, dessert, rice, noodles, meat, seafood, dairy products, egg, soup, fruit, and fried food. The label is represented in integers from 0 to 10.

Data format:

The jpg file name is in the format of '[label]_[id].jpg', for example, 0_5.jpg means this image has an id of 5 and a label of 0.

Output format:

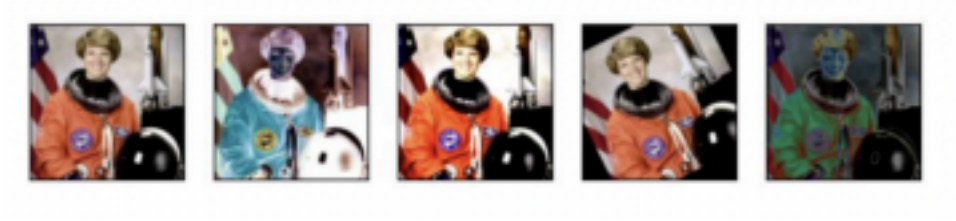
- .csv file
- The submission file includes two columns:
    - Id: 0,1,2,3, etc
    - Label: 0,0,1,5,0, etc

Evaluation metric: Categorization Accuracy

Hints to improve your performance:

1. Tune hyperparameters and change optimizers
2. Design more complex models
3. Cross validation, or resplit train/validation
4. Data augmentation (optional):

a. Try torchvision.transforms to generate more diverse images, for example:



b. Manually modify torch.utils.Dataset to generate the linear combination of two images

i. Change the returned images

ii. Change the label