

## I. Data Preprocessing

- Cifar
  - Random Crop: Randomly Cropping input images into shape (24~32, 24~32)
  - Resize: Resizing images to (32, 32) uniformly
  - Random Flip: Randomly flipping images left or right
  - Rotate: Randomly rotating 20% of the images from -5~5 degrees
  - Random Contrast: Randomly adjusting the contrast(對比度) lower to 1, upper to 1.8
  - Random Saturation: Randomly adjusting the saturation(飽和度) lower to 0.2, upper to 1.8
  - Random Hue: Randomly adjusting the hue(色相) with delta 0.5
  - Translate: Translating 20% of the images in the range of (-10~10, -10~10)
- Cifar - filter
  - 另一個方法是我們先對彩色圖片轉成黑白圖片，並對圖片做前處理，以下是我們所做的處理
    - Ridge detection
    - Sobel
    - Diagonal Sharpen
    - Line Sharpen
  - 除此之外我們也發現在使用模糊化後(blur)反而會造成accuracy下降
  - 同時改變圖片亮度(brightness)也不會提高accuracy
  - 經過實驗發現使用VGG16可以獲得較高accuracy

| models      | val_acc |
|-------------|---------|
| VGG16       | 0.6784  |
| VGG19       | 0.6694  |
| DenseNet201 | 0.6311  |

experiment under same data

| kernel | val_acc |
|--------|---------|
|--------|---------|

|                  |        |
|------------------|--------|
| Blur             | 0.449  |
| Ridge Detection  | 0.6045 |
| Sobel            | 0.6042 |
| Diagonal         | 0.6168 |
| Sharpen          | 0.5809 |
| Line             | 0.5891 |
| all without blur | 0.6314 |

experiment under densent201

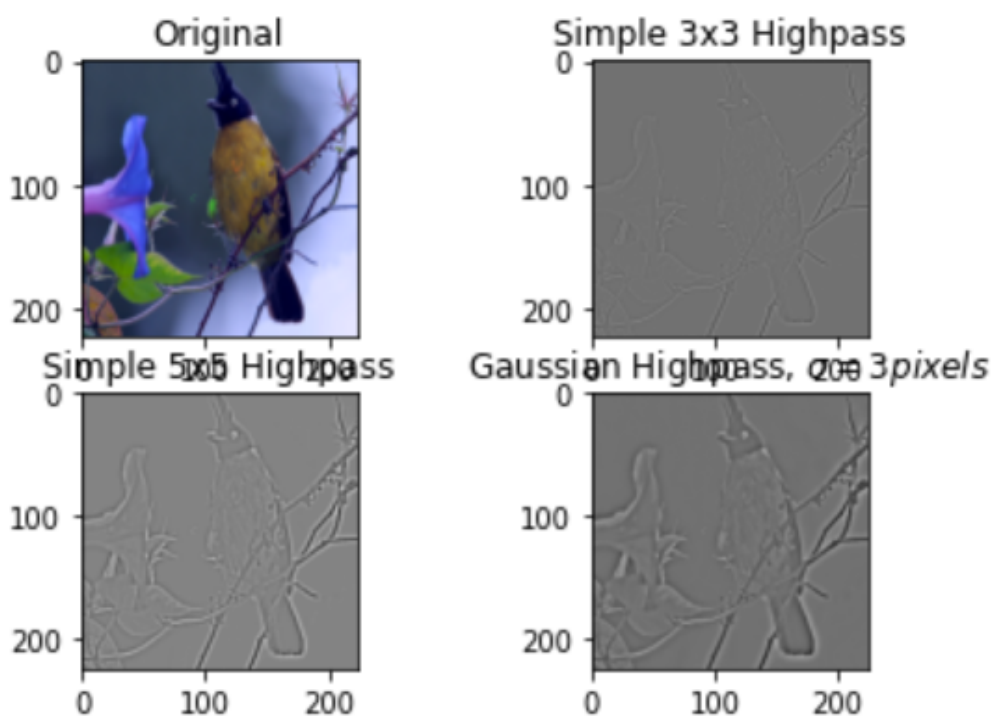
- Imagenet

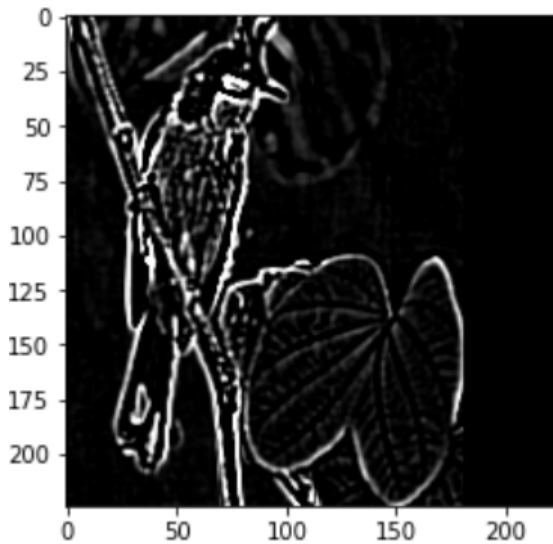
- Random Crop: Randomly Cropping into shape (180~224, 180~224)
- Resize: Resizing images to (224, 224) uniformly
- Random Flip: Randomly flipping images left or right
- Translate: Translating 20% of the images in the range of (-60~60, -60~60)

- Imagenet - filter

- Gray Image
- High Pass filter

一開始的想法是如果noise跟圖片色相有關，不如把圖片灰化並過high pass filter讓他只留下邊緣的線條，結果發現這個方法特別有用





## II. What kind of models you have tried and how did they work

- Cifar

- Resnet-50

- 訓練速度很快, 用來測試 filtering, augmentation 有沒有提高accuracy.
    - testing accuracy: 0.6374

- DenseNet201

- 其中我們將pooling選項設為none, 當時猜想加密資料學到的事應該是讓cnn在max pooling時取到區域中的雜訊, 因此我們在幾次實驗中把pooling設為none。在這個model中我們train了421 epochs, 最後收斂在training accuracy 0.9969, 以及validation accuracy 0.7223。

- VGG-16

- 起初我們也將pooling改為avg pooling或是none, 不過最後仍是使用max pooling有最佳的表現。在這個model我們train了2000 epochstraining accuracy 0.9650, 以及validation accuracy 0.8337。再加上後來的資料前處理, 我們將validation accuracy 0.8701。

- VGG-19

- 因為VGG-19的層數較多, 因此我們預期可以讓features被更準確的抓取, 但是實驗結果也不佳。我們猜想是因為圖片資訊大小是32\*32, 因此沒辦法超越VGG-16的結果。

- VGG-11

- 在使用與VGG19相同的資料訓練後，得到得val\_acc比VGG13和VGG16還差，因此沒有選擇此模型。
- VGG-13
  - 在這個model我們train了227 epochs，在確認收斂後training accuracy 0.8852，以及validation accuracy 0.8227。因為沒有超越遠本VGG-16的表現，因此也沒有再多做實驗。
- ViT
  - Patch size: 8\*8
  - Patches per image: 16



- Patch Encoder: projection(dense to 64) + position embedding(number of patches embedding to 64)
- MLP: [256, 128, 64] Dense layers with dropout rate 0.1
- Validation Accuracy: 0.4971

### III. What problems occurred and how did you solve them

- 我們有嘗試使用encode-decode的模型，但我們發現沒有辦法讓train\_acc, val\_acc一起上升，我們認為可能的原因是透過encode-decode沒有辦法把助教設計的雜訊去除，導致我們所拿到前處理後的圖片無法讓val\_acc上升。
- 有嘗試augmentation加rotation，但表現沒有比較好，角度太高會嚴重降低acc。
- 試過將我們在Imagenet的workflow(將圖片灰階化過high pass filter)套用在cifar10，但由於cifar10解析度太低，如果灰階化再high pass剩下的資訊會太少，train不太起來，準確率約0.5左右