Image Preprocessing

Bert Gollnick

Image Preprocessing

Introduction

- Original image
- Represented by pixels
- Dimensions: (C, H, W)…color, height, width

3 for color image
1 for grayscale

number of
horizontal pixels

number of
vertical pixels

- Pixels are represented as numerical values with [0, …, 255]
- Image conversion required by PyTorch
- Augmentations can improve model performance

Original Image

Image Preprocessing

Resize

- Scales an image (usually down); identical tensor shapes required for all images

Original Image

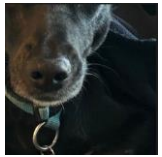
Processed Image

Image Preprocessing

CenterCrop

- crops an image at the center

Original Image



Processed Image

Image Preprocessing

Grayscale

- creates grayscale image, reduces color channel dimension from 3 to 1

Original Image

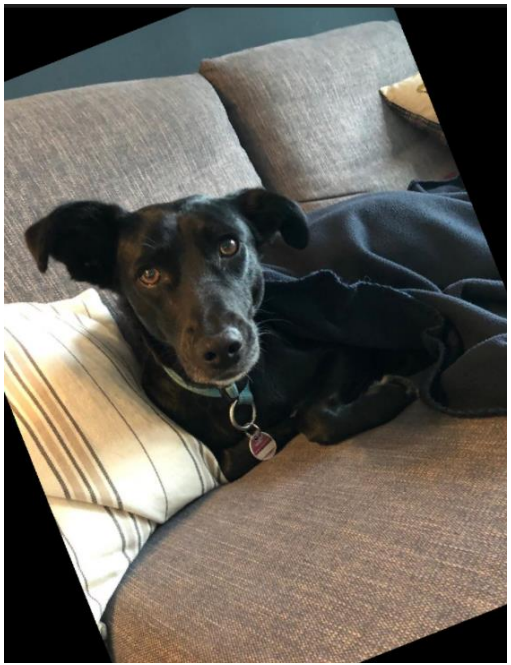
Processed Image

Image Preprocessing

RandomRotation

- randomly rotates image within bounds, usually combined with crop.

Original Image



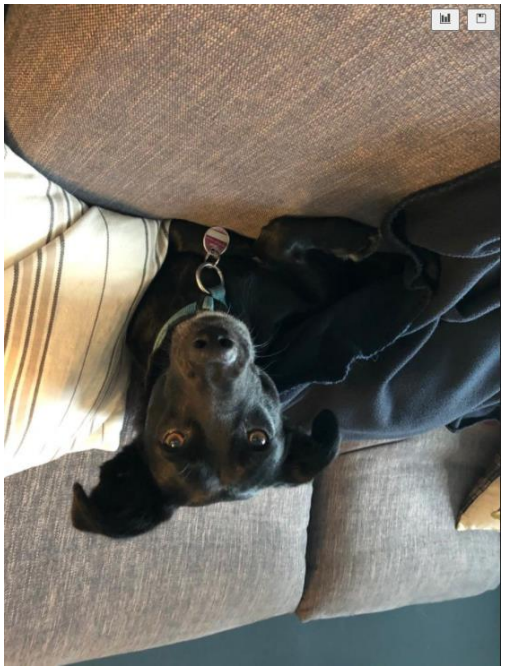
Processed Image

Image Preprocessing

RandomVertialFlip

- flips an image with given probability

Original Image



Processed Image

Image Preprocessing

ToTensor

- converts a PIL image to tensor with dimensions (C, H, W) and value range [0.0, 1.0]

Original Image

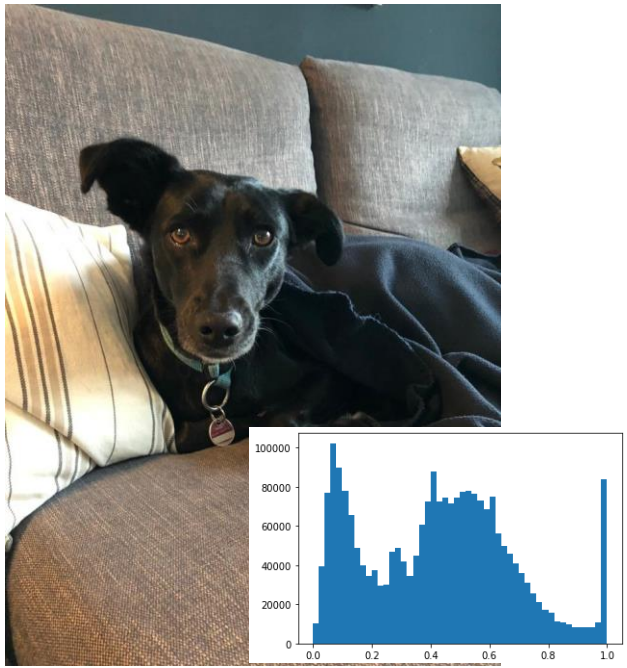
```
tensor([[[0.7647, 0.7882, 0.6235,  ..., 0.2549, 0.2510, 0.2471],
         [0.7412, 0.8000, 0.7255,  ..., 0.2549, 0.2510, 0.2510],
         [0.7098, 0.7725, 0.8118,  ..., 0.2549, 0.2549, 0.2549],
         ...,
         [0.1961, 0.1961, 0.2039,  ..., 0.6510, 0.6078, 0.5765],
         [0.2314, 0.2078, 0.1922,  ..., 0.6078, 0.6078, 0.6392],
         [0.2745, 0.2196, 0.1765,  ..., 0.5725, 0.5725, 0.6392]],

        [[0.6902, 0.7137, 0.5490,  ..., 0.3686, 0.3647, 0.3608],
         [0.6667, 0.7255, 0.6510,  ..., 0.3686, 0.3647, 0.3647],
         [0.6353, 0.6980, 0.7373,  ..., 0.3686, 0.3686, 0.3686],
         ...,
         [0.1765, 0.1765, 0.1843,  ..., 0.5961, 0.5529, 0.5216],
         [0.2118, 0.1882, 0.1725,  ..., 0.5608, 0.5608, 0.5922],
         [0.2549, 0.2000, 0.1569,  ..., 0.5255, 0.5255, 0.5922]],

        [[0.6235, 0.6471, 0.4824,  ..., 0.3922, 0.3882, 0.3843],
         [0.6000, 0.6588, 0.5843,  ..., 0.3922, 0.3882, 0.3882],
         [0.5686, 0.6314, 0.6706,  ..., 0.3922, 0.3922, 0.3922],
         ...,
         [0.1529, 0.1529, 0.1608,  ..., 0.5451, 0.5020, 0.4706],
         [0.1882, 0.1647, 0.1490,  ..., 0.5059, 0.5059, 0.5373],
         [0.2314, 0.1765, 0.1333,  ..., 0.4706, 0.4706, 0.5373]]])
```

Processed Image

Image Preprocessing

Normalize

- normalizes an image with mean and standard deviation



Original Image (created from tensor)

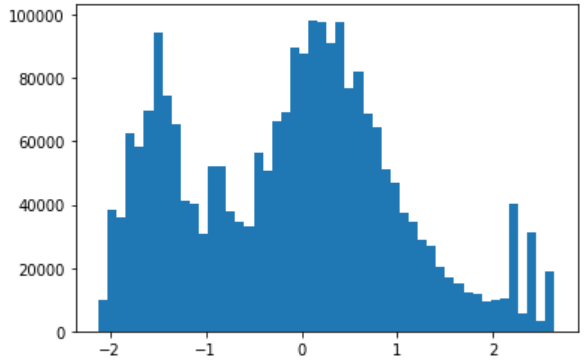
Batch of images

e.g. ImageNet

mean = (0.485, 0.456, 0.406)

Sd = (0.229, 0.224, 0.225)

Image is centered around given mean and has given standard deviation.



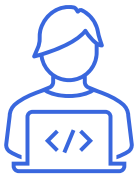
Processed Image

Image Preprocessing

Compose

- Performs several transformations at once
- composed steps are applied to all images

```
preprocess_steps = transforms.Compose([
    transforms.Resize(300),  # better (300, 300)
    transforms.RandomRotation(50),
    transforms.CenterCrop(500),
    transforms.Grayscale(),
    transforms.RandomVerticalFlip(),
    transforms.ToTensor(),
    transforms.Normalize((0.485, 0.456, 0.406), (0.229, 0.224, 0.225)),
])
x = preprocess_steps(img)
```



Let‘s find out in coding lecture.