

25MML0014
ANUSREE V

EDGE INTELLIGENCE

1.loading problem

The main problem faced is loading dataset because it contains large mobs. So I just downloaded in terminal .it took 28 minutes for loading as it is huge data.

So I just compiled and tried everything in collab first , Jupyter notebook it is taking hours to loading the dataset.

```
[*]: import urllib.request
import tarfile
import os
from PIL import Image
url = "http://vision.stanford.edu/Aditya86/ImageNetDogs/images.tar"
tar_path = "dogs.tar"

print("Downloading..")
urllib.request.urlre
# Download the dataset tar file
curl -L -o dogs.tar "http://vision.stanford.edu/Aditya86/ImageNetDogs/images.tar"
print("Extracting...")
with tarfile.open(tar_path) as tar:
    tar.extractall()
DATASET_DIR = "dogs/"

shapes = sorted([
    Image.open(os.path.join(DATASET_DIR, root, file))
    for root, _, files in os.walk(DATASET_DIR)
    for file in files
    if file.lower().endswith(('.jpg', '.png', '.jpeg'))])

print("Total images:", len(shapes))
print("Unique shapes:", len(set(shapes)))

Downloading...
```

```
for root, _, files in os.walk(DATASET_DIR):
    for file in files:
        if file.lower().endswith(('.jpg', '.png', '.jpeg')):

            img_path = os.path.join(root, file)

            img = Image.open(img_path)
            shape = img.size
            all_shapes.append(shape)
            count += 1
            print(f"Image {count}: {shape}")

print("\nTotal images found:", count)
print("Unique image shapes:", set(all_shapes))

...
-----
FileNotFoundError                         Traceback (most recent call last)
/tmp/ipython-input-2013667079.py in <cell line: 0>()
      8         img_path = os.path.join(root, file)
      9
--> 10     img = Image.open(img_path)
     11     shape = img.size
     12     all_shapes.append(shape)

/usr/local/lib/python3.12/dist-packages/PIL/Image.py in open(fp, mode, formats)
 3511     if is_path(fp):
 3512         filename = os.fspath(fp)
-> 3513         fp = builtins.open(filename, "rb")
 3514         exclusive_fp = True
 3515     else:
```

img_path is given wrong

```
all_shapes = []
count = 0

for root, _, files in os.walk(DATASET_DIR):
    for file in files:
        if file.lower().endswith((".jpg", ".png", ".jpeg")):

            img_path = os.path.join(root, file)

            img = Image.open(img_path)
            shape = img.size
            all_shapes.append(shape)
            count += 1
            print(f"Image {count}: {shape}")

print("\nTotal images found:", count)
print("Unique image shapes:", set(all_shapes))
```

```
import numpy as np
image_list = []
shape_list = []

limit = 25
count = 0

for root, _, files in os.walk(DATASET_DIR):
    for file in files:
        if file.lower().endswith((".png")):
            img_path = os.path.join(root, file)
            img = Image.open(img_path).convert("RGB")

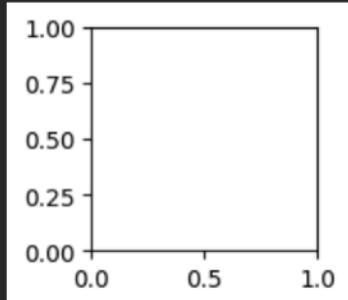
            image_list.append(np.array(img))
            shape_list.append(img.size)

            count += 1
            if count == limit:
                break
    if count == limit:
        break

plt.figure(figsize=(10, 10))
for i in range(limit):
    plt.subplot(5, 5, i+1)
    plt.imshow(image_list[i])
    plt.title(str(shape_list[i]))
    plt.axis("off")

plt.suptitle("First 25 Images & Their Shapes", fontsize=16)
plt.tight_layout()
plt.show()
```

```
...-----  
IndexError Traceback (most recent call last)  
/tmp/ipython-input-3648484427.py in <cell line: 0>()  
    25     for i in range(limit):  
    26         plt.subplot(5, 5, i+1)  
----> 27         plt.imshow(image_list[i])  
    28         plt.title(str(shape_list[i]))  
    29         plt.axis("off")  
  
IndexError: list index out of range
```



Added jpg, jpeg and corrected this error

```
if file.lower().endswith((".jpg", ".png", ".jpeg")):
```

```
import os  
from PIL import Image  
import numpy as np  
  
DATASET_DIR = "dogs/Images/"  
  
TARGET_SIZE = (128, 128)  
  
resized_images = []  
shapes = []  
count = 0  
  
print("Resizing all images to:", TARGET_SIZE)  
  
for root, _, files in os.walk(DATASET_DIR):  
    for file in files:  
        if file.lower().endswith((".jpg", ".jpeg", ".png")):  
            img_path = os.path.join(root, file)  
            img = Image.open(img_path).convert("RGB")  
            img = img.resize(TARGET_SIZE)  
            resized_images.append(img)  
            shapes.append(img.shape)  
  
        count += 1  
        print(f"Image {count} resized to {img.shape}")  
  
resized_images = np.array(resized_images)  
  
print("\nTotal images resized:", resized_images.shape[0])  
print("Each image final shape:", resized_images[0].shape)
```

```
resized_images = np.array(resized_images)

print("\nTotal images resized:", resized_images.shape[0])
print("Each image final shape:", resized_images[0].shape)

... Resizing all images to: (128, 128)
-----
AttributeError Traceback (most recent call last)
/tmp/ipython-input-3203422548.py in <cell line: 0>()
     28
     29         resized_images.append(img)
--> 30     shapes.append(img.shape)
     31
     32     count += 1

AttributeError: 'Image' object has no attribute 'shape'
```

Converted the image into numpy array and corrected this

```
for root, _, files in os.walk(DATASET_DIR):
    for file in files:
        if file.lower().endswith((".jpg", ".jpeg", ".png")):
            img_path = os.path.join(root, file)

            img = Image.open(img_path).convert("RGB")

            img = img.resize(TARGET_SIZE)

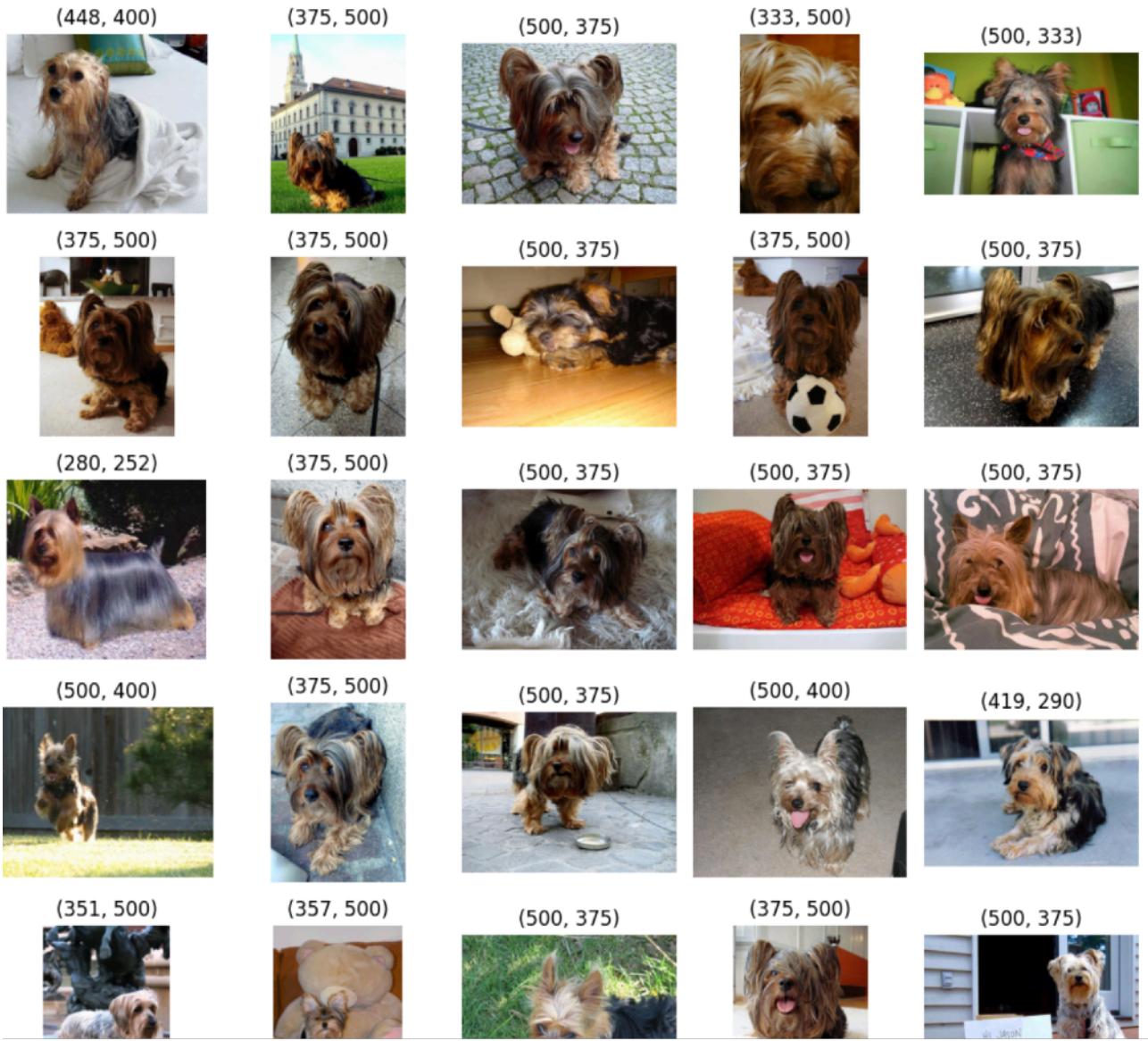
            img_np = np.array(img)

            resized_images.append(img_np)
            shapes.append(img_np.shape)

            count += 1
            print(f"Image {count} resized to {img_np.shape}")

resized_images = np.array(resized_images)
```

First 25 Images & Their Shapes



Resized Images (128x128)

