DS Creator

November 7, 2022

[ ]:
```
from ipywidgets import IntProgress
from IPython.display import display

import tensorflow as tf
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

import kaggle
import os
import glob
import shutil
print("Num GPUs Available: ", len(tf.config.experimental.
 ↪list_physical_devices('GPU')))


BASE_DIR = os.getcwd()
```

[ ]:
```
# Download CelebA Dataset from Kaggle
!kaggle  datasets download -d jessicali9530/celeba-dataset
!unzip -d dataset -u -q celeba-dataset.zip
```

[ ]:
```
#Sort images into categories
df = pd.read_csv('dataset/list_attr_celeba.csv')
df.head()
is_male_df = df["Male"]

ds_PATH = "dataset/img_align_celeba/img_align_celeba/"

for file in glob.glob(os.path.join(ds_PATH, "*.jpg")):
    file_number = file.strip(ds_PATH).strip(".jpg")
    if is_male_df[int(file_number)-1] == 1:
        shutil.copy2(file, os.path.join(BASE_DIR, "dataset/preprocessed/male",␣
 ↪file_number+".jpg"))
    else:
        shutil.copy2(file, os.path.join(BASE_DIR, "dataset/preprocessed/
 ↪female", file_number+".jpg"))
```

[ ]:
```
IMAGE_WIDTH = 64
IMAGE_HEIGHT = 64
BATCH_SIZE = 128
```

[ ]:
```
# generate and test dataset
AUTOTUNE = tf.data.AUTOTUNE
train_ds = tf.keras.utils.image_dataset_from_directory(
    os.path.join(BASE_DIR, "dataset/preprocessed"),
    image_size=(IMAGE_HEIGHT, IMAGE_WIDTH),
    batch_size=BATCH_SIZE)
train_ds.cache().prefetch(buffer_size=AUTOTUNE)
```

[ ]:
```
# plot example images
class_names = train_ds.class_names
plt.figure(figsize=(10, 10))
for images, labels in train_ds.take(1):
    for i in range(9):
        ax = plt.subplot(3, 3, i + 1)
        plt.imshow(images[i].numpy().astype("uint8"))
        plt.title(class_names[labels[i]])
        plt.axis("off")
```