data-acquisition.md 12/14/2022

Transfer Learning for Custom Datasets in the Small-Data Regime for Basement

Milestone 2: Data Acquisition

For the basement I shortlisted following 10 product catogries that HD sells and are visible in their video:

- 1. Pillar Wood/Steel
- 2. Television
- 3. Photo Frame
- 4. Lamp and Lighting
- 5. Fireplace
- 6. Flooring
- 7. Sofa
- 8. Table
- 9. Installation Service
- 10. Window

I used Google Custom Search API to get images for disered labels. Here is the code for a particular lable:

```
# !pip install google-api-python-client
import requests
from PIL import Image
import os
from apiclient.discovery import build
resource = build("customsearch", 'v1', developerKey=api key).cse()
items = []
search_string = "basement wire"
cx = "xxxxxxxxxxx" #Custom Search Engine ID
for rng in range(1,100,10):
   print(rng, end=' ')
   result = resource.list(q=search_string, cx='c027b0cbdfed84f0b', searchType =
'image',start=rng).execute()
   for item in result['items']:
       items.append([item['title'], item['link']])
# Saving Images into folder
dir name = './wire data/'
if dir_name not in os.listdir():
   os.makedirs('dir_name')
i=0
for item in items:
   img url = item[1]
   response = requests.get(img_url)
```

data-acquisition.md 12/14/2022

```
if response.status_code:
    fp = open('./wire/+'+str(i)+'.png', 'wb')
    fp.write(response.content)
    fp.close()
print(i,item[1])
i+=1
```