paperdl

Release 0.1.0

Charles

Jul 20, 2022

CONTENTS

STATEMENTS

This repo is created for learning python.

If I find that anyone leverage this project in an illegal way, I will delete this project immediately.

INSTALL PAPERDL

2.1 Environment

- OS: Linux or macOS or Windows
- Python version: Python3.6+

2.2 Pip install

You can run the following command in the terminal to install paperdl (make sure that the python is in the environment variable):

```
pip install paperdl --upgrade
```

2.3 Source code install

1.Online

You can run the following command in the terminal to install paperdl (make sure that the python and git are in the environment variable):

```
pip install git+https://github.com/CharlesPikachu/paperdl.git@master
```

2.Offline

You can first run the following command in the terminal to download the source code in your computer:

```
git clone https://github.com/CharlesPikachu/paperdl.git
```

Then, enter the corresponding directory:

```
cd paperdl
```

Finally, run the following command in the terminal to install paperdl:

```
python setup.py install
```

QUICK START

3.1 Calling API

If you want to search and download papers from arxiv and google scholar, you can write codes as follow:

```python
from paperdl import paperdl

config = {'logfilepath': 'paperdl.log', 'savedir': 'papers', 'search_size_per_source': 5,
→ 'proxies': {}}
target_srcs = ['arxiv', 'googlescholar']
client = paperdl.Paperdl(config=config)
client.run(target_srcs)
```

In addition, if you can not visit google, you can set config as follow:

```python
config = {'logfilepath': 'paperdl.log', 'savedir': 'papers', 'search_size_per_source': 5,
→ 'proxies': {}, 'area': 'CN'}
```

You can also only download papers by using sci-hub as follow:

```python
from paperdl import paperdl

config = {'logfilepath': 'paperdl.log', 'savedir': 'papers', 'search_size_per_source': 5,
→ 'proxies': {}}
client = paperdl.SciHub(config=config, logger_handle=paperdl.Logger('paper.log'))
paperinfo = {
    'savename': '9193963',
    'ext': 'pdf',
    'savedir': 'outputs',
    'input': 'https://ieeexplore.ieee.org/document/9193963/',
    'source': 'scihub',
}
client.download([paperinfo])
```

Here is a screenshot:

3.2 Calling EXE

You can directly leverage paperdl in the terminal, and the usage is as follow:

```
Usage: paperdl [OPTIONS]

Options:
  --version             Show the version and exit.
  -m, --mode TEXT       the used mode, support "search" and "download"
  -i, --inp TEXT        the paper to download, the supported format is the
                        same as sci-hub
  -s, --source TEXT     the used source, support "arxiv", "scihub" and
                        "googlescholar", you can use "," to split multi
                        sources
  -d, --savedir TEXT    the directory for saving papers
  -l, --logfilepath TEXT  the logging filepath
  -z, --size INTEGER    search size per source
  -p, --proxies TEXT    the proxies to be adopted
  -a, --area TEXT       your area, support "CN" and "EN"
  -c, --cookie TEXT     the cookie copied from the target website, only used
                        in "baiduwenku"
  --help                Show this message and exit.
```

Here is an example:

```
paperdl -i https://ieeexplore.ieee.org/document/7485869/ -m download
```

FOUR

CHANGELOG

2022-02-23

- Version: v0.1.3,

- Update: support arxiv, googlescholar and scihub.

2022-02-25

- Version: v0.1.4,

- Update: some improvements, like adding exe auto building and the progress bar.

2022-05-08

- Version: v0.1.5,

- Update: support baiduwenku.

2022-05-13

- Version: v0.1.6,

- Update: fix the bugs in downloader.

RECOMMENDED PROJECTS

- Games: Create interesting games by pure python.

- DecryptLogin: APIs for loginning some websites by using requests.

- Musicdl: A lightweight music downloader written by pure python.

- Videodl: A lightweight video downloader written by pure python.

- Pytools: Some useful tools written by pure python.

- PikachuWeChat: Play WeChat with itchat-uos.

- SSSegmentation: An Open Source Strongly Supervised Semantic Segmentation Toolbox Based on PyTorch.

- Pydrawing: Beautify your image or video.

- ImageCompressor: Image compressors written by pure python.

- FreeProxy: Collecting free proxies from internet.

- Constellation: Beautiful starry sky written by js.

- Paperdl: Search and download paper from specific websites.

- Sciogovterminal: Browse "The State Council Information Office of the People's Republic of China" in the terminal.

- CodeFree: Make no code a reality.

- DeepLearningToys: Some deep learning toys implemented in pytorch.

- DataAnalysis: Some data analysis projects in charles_pikachu.

- Imagedl: Search and download images from specific websites.

- Pytoydl: A toy deep learning framework built upon numpy.

- NovelDL: Search and download novels from some specific websites.

ABOUT ME

I'm a student whose research interests include computer vision and information security.

WeChat public account: Charles_pikachu

Github: https://github.com/CharlesPikachu

Zhihu: https://www.zhihu.com/people/charles_pikachu

Bilibili: https://space.bilibili.com/406756145

Email: charlesblwx@gmail.com