



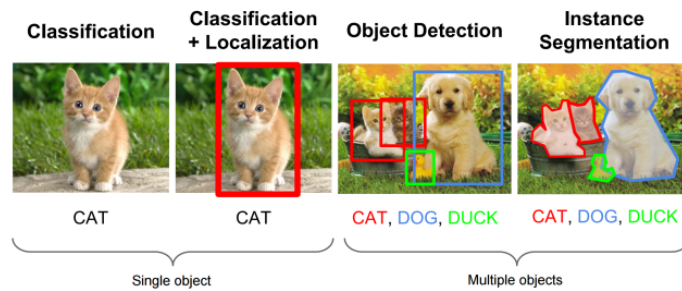
A simple way to collect your deep learning image dataset



Catch Zeng

[Follow](#)

Feb 24 · 2 min read



Deep Learning has become the go-to method for solving many challenging problems. As we know, with enough training, a deep network can segment and identify the “key points” in the image.

If a very simple mechanism is large enough, it will have a magical effect.

Therefore, this well-functioning deep learning **requires a lot of data**. The more training data, the better the accuracy of the model.

But where do we get all this data from? Well-annotated data can be both **expensive** and **time-consuming** to acquire. Hiring people to manually collect images and label them is **not efficient** at all. And, in the deep learning era, **data is very well arguably your most valuable resource**.

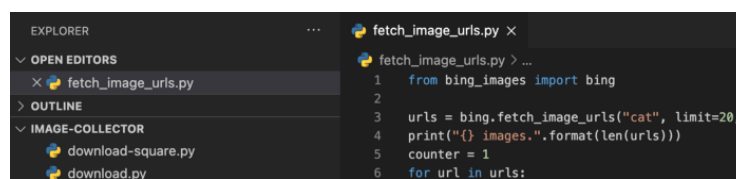
Here, I show a simple way to collect your deep learning image dataset.

The [bing-images](#) is a Python library to **fetch image URLs** and download using **multithreading** from [Bing.com](#). It has the following features

- Support **file type** filters.
- Support [Bing.com](#) *filterui* filters.
- Download using **multithreading** and custom thread pool size.
- Support **purely** obtaining the image URLs.

Demo

Create a demo project, called *image-collector* here.



```
fetch_image_urls.py 7 | print('{}: {}'.format(counter, url))
8 | counter += 1
```

Install bing-images

Requirements

- Install Google Chrome Browser.
- Download `chromedriver` from here.
- Add `chromedriver` to PATH.

Fetch image URLs

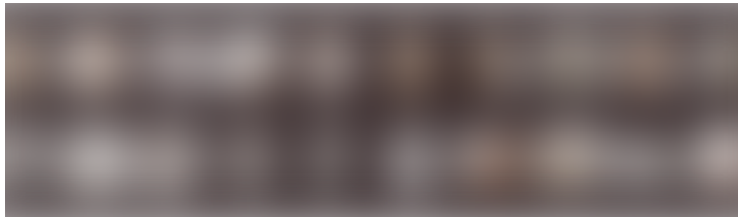
`fetch_image_urls.py`

Run

Download using multithreading

download.py

Run



Download square black-white images

`download-square.py`




The detailed code is at https://github.com/CatchZeng/bing_images.
See you!

Sign up for Analytics Vidhya News Bytes

By Analytics Vidhya

Latest news from Analytics Vidhya on our Hackathons and some of our best articles! [Take a look.](#)

Your email

 Get this newsletter

By signing up, you will create a Medium account if you don't already have one. Review our [Privacy Policy](#) for more information about our privacy practices.

Deep Learning

Image Collection

Bing Image Search

Bulk Image Downloader

Image Scraping



WRITTEN BY

Catch Zeng

Deep Learning and DevOps enthusiast

Follow



Analytics Vidhya

Analytics Vidhya is a community of Analytics and Data Science professionals. We are building the next-gen data science ecosystem <https://www.analyticsvidhya.com>

Follow

More From Medium

How To Train a Core ML Model on Your Device

Anupam Chugh in Better Programming



Building a Multiple Object Detection Model with TensorFlow's Object Detection API

Ronak Bhatia



Of Gradients and Matrices

Paolo Perrotta



Review: ResNet-38—Wider or Deeper ResNet? (Image Classification & Semantic Segmentation)

Sik-Ho Tsang



Get Started -Model Building and Machine Learning Pipeline with Azure ML Studio For Free



Combinatorial Optimization: from theory to code using Google's OR tools



Gold Price Prediction using Machine Learning with Python

Mahmoud Galib Mahmoud in



How Perceptrons solve the linearly separable problems

Prakash Kedia in ML applications



Learn more.

Medium is an open platform where 170 million readers come to find insightful and dynamic thinking. Here, expert and undiscovered voices alike dive into the heart of any topic and bring new ideas to the surface. [Learn more](#)

Make Medium yours.

Follow the writers, publications, and topics that matter to you, and you'll see them on your homepage and in your inbox. [Explore](#)

Write a story on Medium.

If you have a story to tell, knowledge to share, or a perspective to offer — welcome home. It's easy and free to post your thinking on any topic. [Start a blog](#)