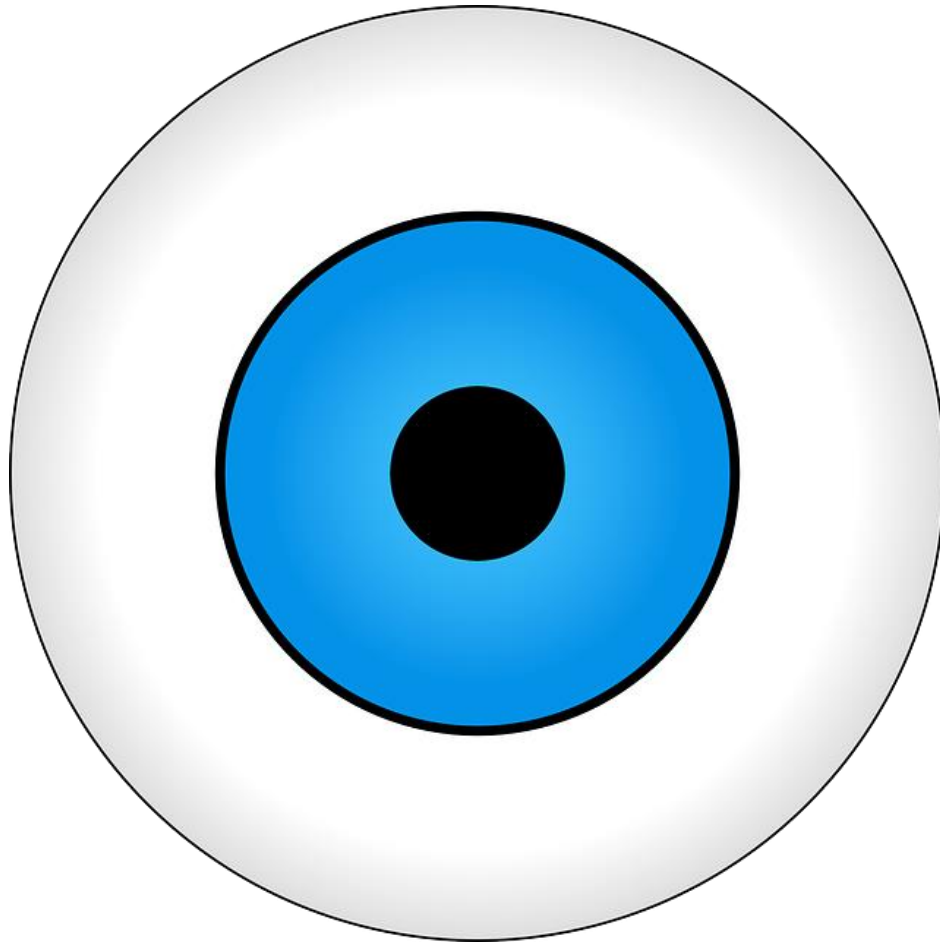# Computer Vision Resources

**Satya Mallick, Ph.D.**

LearnOpenCV.com

Author's Note

Congratulations! By downloading this resource guide, you have embarked on a journey of learning. This guide is not a laundry list of all available computer vision resources. On the contrary, it is a curated list of things I find useful in my work. It is often wise to leave out ingredients from a recipe to improve it, and so I have decided to leave out resources that may overwhelm a beginner. However, if you do have a resource that you find useful, please email me at spmallick@learnopencv.com

Who is this guide for?

This guide is for programmers, hackers, engineers, scientists, students, and self-starters. It is for those creative people who have an itch to learn something new and build something useful and beautiful. It is for people who take pride in their work and are craftsmen at heart. It is for men and women who believe in sharpening their tools and improving their craft on a regular basis. It is for those who believe that learning is a continuous process, and that there are smart ways to learn fast. It is for tinkerers who can learn by reading but prefer to learn by doing. Lastly, it is for people who invest in themselves by learning something new every day and are eager to contribute back to the community to enrich others!

Books
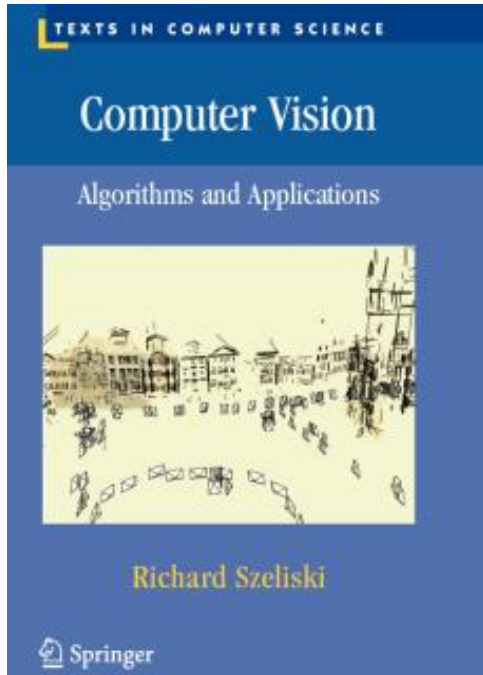
1. Computer Vision: Algorithms and Applications

Author: Richard Szeliski

Summary: This is by far the best book on computer vision I have used. This is a good basic reference book for a wide variety of computer vision topics — image formation, image processing, feature detection and matching, image segmentation, image alignment, structure from motion, motion estimation, image stitching, computational photography, stereo correspondence, 3D reconstruction, image based rendering, and recognition.

Free version: http://szeliski.org/Book/
**Buy at Amazon (Hardcover)**
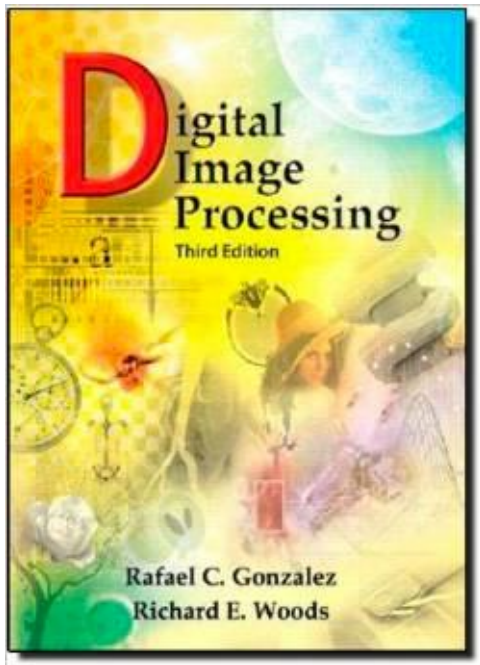
2. Digital Image Processing

Authors: Rafael C. Gonzalez and Richard E. Woods
Summary: This is a good introductory book in image processing. It may be challenging to understand Computer Vision literature without a basic understanding of image processing concepts. The book covers digital image fundamentals, image enhancement (spatial and frequency domains), image restoration, color image processing, wavelets and multi-resolution processing, image compression, morphological operations, segmentation, and a bit of object detection.

Buy at Amazon (Hardcover and Paperback)
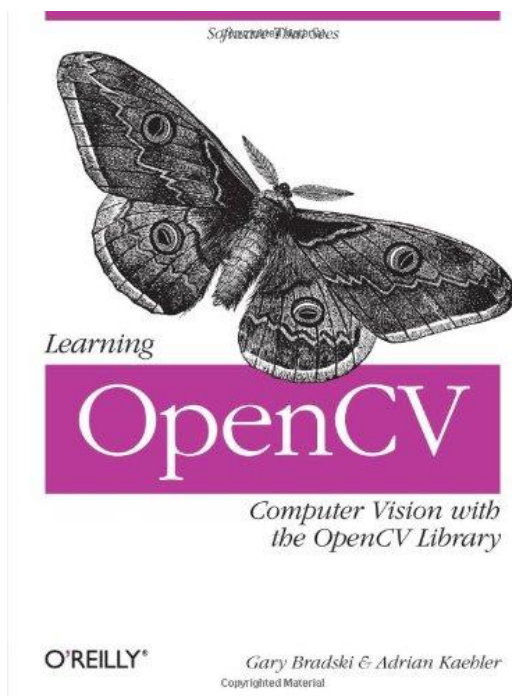
3. Learning OpenCV

Authors: Gary Bradski & Adrian Kaehler
Summary: Gray Bradski started OpenCV and this book is a great introductory book for learning OpenCV. The book comes with links to code samples and tutorials. The only downside is that this version of the book does not cover OpenCV 3.

Note: The examples in the book are in C++ only.

Buy at Amazon

4. Practical Python and OpenCV

Authors: Adrian Rosebrock
Summary: Great introductory book for learning OpenCV using Python. It contains example code, and interesting case studies. This digital book comes with a 30-day money back guarantee, so it is risk free! You also receive free updates to the book as it is revised (e.g. when OpenCV 3 is released some of the code will be revised).

Buy at PyImageSearch.com

Python Books

1.   Fluent Python

Authors: Luciano Ramalho

Summary: Great book on effective, idiomatic Python. However, it is not a basic book - if you do not know the Python language at all, it will be too complex for you
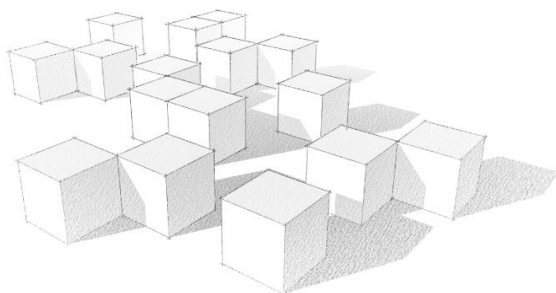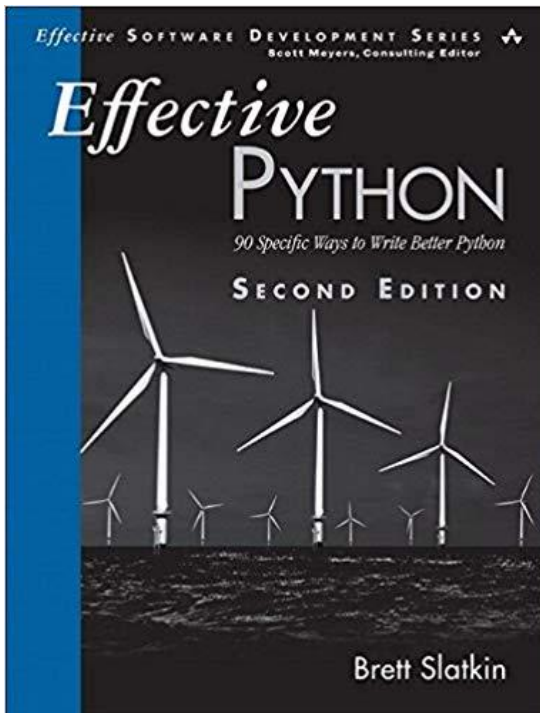
Book at O'Reilly store

2.   From Python to Numpy

Authors: Nicolas P. Rougier

Summary: If you already know Python and need to use the NumPy superpower, this book is for you! Nicolas has done the great job describing the tips and tricks of NumPy starting from the basics and up to the mastery.

Link to the free copy of the book

3.   Effective Python

Author: Brett Slatkin

Summary: This second edition book provides insight into the "Pythonic" way of writing programs: the best way to use Python. The author brings together 90 Python best practices, tips, and shortcuts, and explains them with realistic code examples so that you can embrace Python with confidence.

Note: All of the examples from the book could be found in **the github repo.**

Buy at Amazon
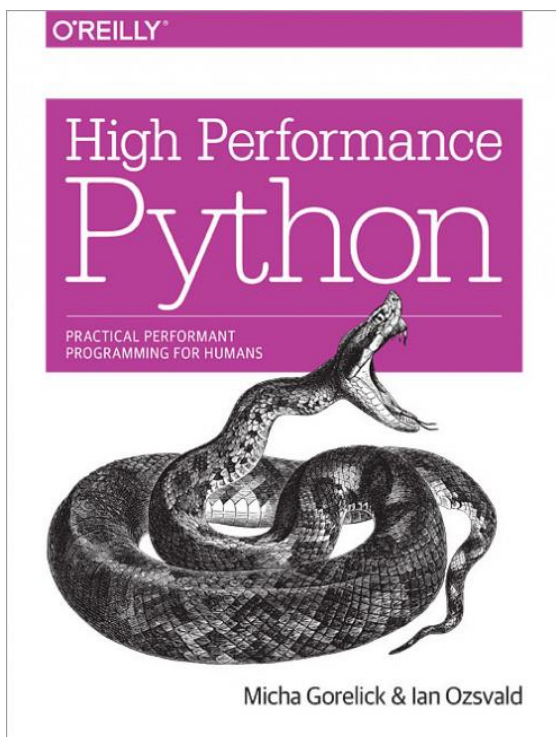
4.   High Performance Python

Author: Ian Ozsvald, Micha Gorelick

Summary: By exploring the fundamental theory behind design choices, this practical guide helps you gain a deeper understanding of Python's implementation. You will learn how to locate performance bottlenecks and significantly speed up your code in high-data-volume programs.

Book at O'Reilly store

PyTorch books
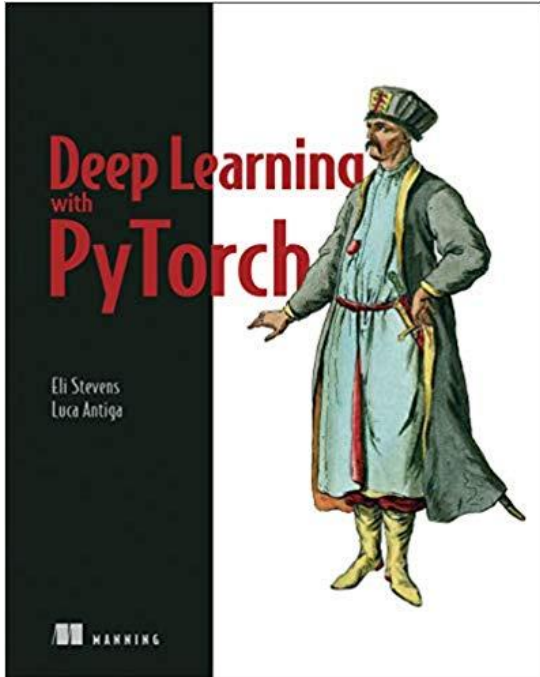
1.   Deep Learning with PyTorch

Authors: Eli Stevens, Luca Antiga

Summary: Deep Learning with PyTorch provides a detailed, hands-on introduction to building and training neural networks with PyTorch. It includes:

- Introduction to deep learning and the PyTorch library
- Pre-trained networks
- Tensors
- The mechanics of learning
- Using a neural network to fit data

Link to the free copy of the book

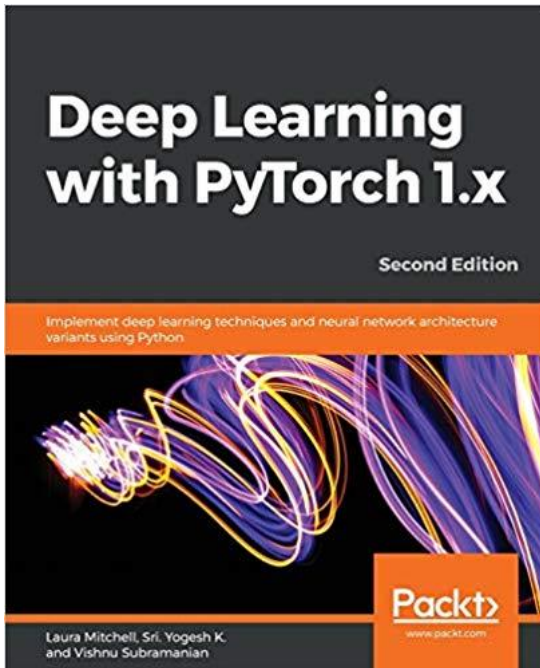## 2. Deep Learning with PyTorch 1.x: Implement deep learning techniques and neural network architecture variants using Python, 2nd Edition

Authors: Laura Mitchell, Sri. Yogesh K., Vishnu Subramanian

Summary: Freshly published in November 2019, this book provides the fundamental aspects that power modern deep learning and covers the new features of the PyTorch 1.x library.

Buy at Amazon

DL books
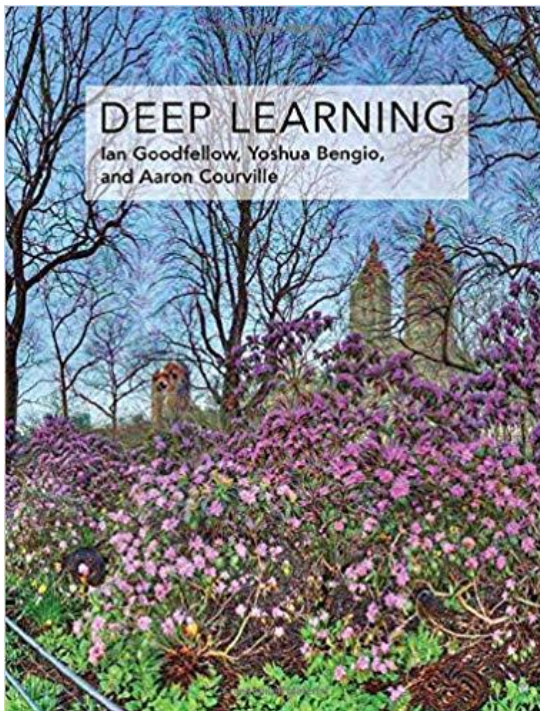
1. Deep Learning

Authors: Ian Goodfellow, Yoshua Bengio, Aaron Courville.

Summary: This book is widely regarded as Deep Learning's "Bible". This is probably the most detailed book you can read on Deep Learning and Neural Nets, written by three experts, including one of the field's godfathers. The book is quite technical, but in describing everything you need to learn to get going, the writers did a great job.

Free online version
Buy at Amazon
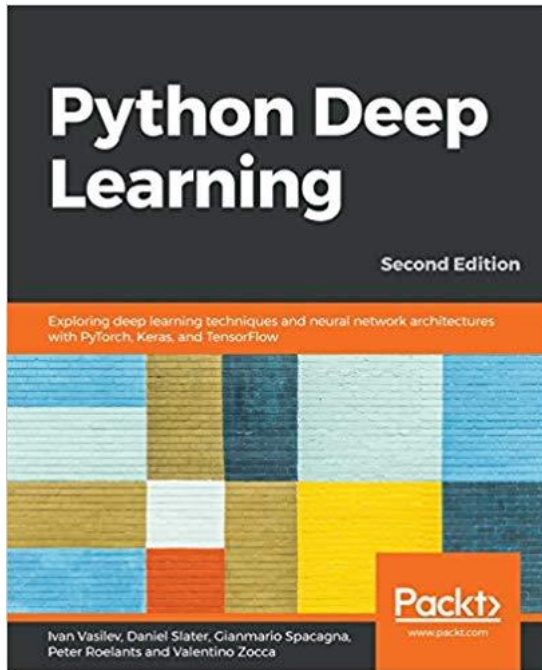
2. Python Deep Learning: Exploring deep learning techniques and neural network architectures with PyTorch, Keras, and TensorFlow

Authors: Ivan Vasilev, Daniel Slater, Gianmario Spacagna, Peter Roelants, Valentino Zocca

Summary: This book will get you up to speed with deep learning, deep neural networks, and how to train them with the most popular Python frameworks. It provides not only in-detail explanation of the basics of neural networks, but gradually introduces more complex models like recurrent networks and GANs in an easy to understand manner.

Buy at Amazon

Online resources

- https://distill.pub/ - cool visualizations

- https://www.youtube.com/watch?v=aircAruvnKk - youtube channel of 3 Blue 1 Brown with neural networks explained by visual examples

- https://www.youtube.com/user/keeroyz/videos?sort=p&view=0&flow=grid - Two Minute Papers' Channel provides an interesting and shallow overview of some deep learning research papers

- http://www.arxiv-sanity.com/ - aggregator of papers from arXiv.com that allows researchers to keep track of recent or popular papers on different themes of Machine Learning

- https://pytorch.org/tutorials/ - official tutorials on using PyTorch from the creators

- https://github.com/vinhkhuc/PyTorch-Mini-Tutorials - mini-tutorials on PyTorch to get to know the framework

- https://github.com/utkuozbulak/pytorch-custom-dataset-examples - some custom dataset examples for PyTorch

- https://github.com/stared/thinking-in-tensors-writing-in-pytorch - a hands-on deep learning intro with some explanation of the mathematics behind

Software & Libraries

OpenCV (http://opencv.org/)

Summary

    The biggest and the most extensive open source computer vision library. OpenCV has more than 47 thousand people of user community and estimated number of downloads exceeding 10 million.

Languages

    C/C++ with interfaces to Python and Java.

Platforms

    Windows, Linux, Mac OS, iOS, Android, Raspberry Pi, and NVIDIA Jetson TK1.

License (http://opencv.org/license.html)

    BSD: It is free for both academic and commercial use.

    Note: Not all parts of OpenCV are free.

VLFeat (http://www.vlfeat.org/)

Summary

    Computer vision algorithms specializing in image understanding and local features extraction and matching.

Languages

    C with interfaces in MATLAB

Platforms

    Windows, Mac OS X, and Linux.

License

    BSD: It is free for academic and commercial use.

SimpleCV (http://simplecv.org/)

Summary

    SimpleCV is an open source wrapper around computer vision libraries such as OpenCV that hides some of its complexities.

Languages

    Python

 Platforms

Windows, Mac OS X, Linux, and Raspberry Pi.

License

BSD: It is free for academic and commercial use.

MATLAB CV Toolbox

((http://www.mathworks.com/products/computer-vision/)

Summary

A computer vision toolbox for MATLAB.

Languages

MATLAB

Platforms

Windows, Mac OS X and Linux.

License

MATLAB's license. Requires Image Processing Toolbox. The total cost of installing MATLAB ($2,150) + Image Processing Toolbox ($1, 000) + Computer Vision Toolbox ($1,350) = $4500. Student licenses are much cheaper though (few hundred dollars).

Python Libraries

One of the main advantages of using OpenCV with Python is the vast number of scientific libraries available for Python. Here are a few libraries you will find useful. The first three libraries — NumPy, SciPy and Matplotlib — are part of the SciPy stack. When used together, they pretty much replace MATLAB.

1. **NumPy (http://www.numpy.org)** : NumPy adds support for large, multi-dimensional arrays and matrices to Python. It also consists of a large library of high-level mathematical functions to operate on these arrays. OpenCV images are read in as NumPy arrays. Several other math, image processing, and machine learning libraries are built on top of NumPy.

2. **SciPy** (**http://scipy.org/scipylib/index.html**) : SciPy is a powerful scientific library built on top of NumPy. It's sub packages include linalg (linear algebra), optimize (optimization and root-finding routines), stats (statistical distributions and functions), ndimage (N-dimensional image processing), interpolate (interpolation and smoothing splines) , fftpack   (Fast Fourier Transform routines), cluster (Clustering algorithms) and many more.

3. **matplotlib** (**http://matplotlib.org/**) : An excellent 2D plotting library for Python that is every bit as powerful as MATLAB. You can generate plots, histograms, power spectra, bar charts, scatterplots, etc, with just a few lines of code.

4. **scikit-learn** (**http://scikit-learn.org/**) : As a computer vision programmer / engineer, you will inevitably need a good machine learning library and scikit-learn serves that purpose well. It uses numpy/scipy idioms and provides algorithms for preprocessing data, classification, regression, clustering, dimensionality reduction, and model selection.

PyTorch Libraries

Here is a list of handy libraries to help you with developing under PyTorch framework.

1. **Albumentations (https://albumentations.readthedocs.io/)** : A fast image augmentation library based on highly-optimized OpenCV library. Super simple yet powerful interface for different tasks (like segmentation, detection, etc.), that could be easily customized. Even though the library could be used not only under PyTorch framework, it has PyTorch helpers in its submodule **albumentations.pytorch** to make working with PyTorch tensors as easy as possible.

2. **Tensorboard (https://www.tensorflow.org/tensorboard/)** : Starting with version 1.2 PyTorch integrates with TensorBoard, a tool designed for visualizing the results of neural network training runs. It comes extremely helpful to see what's happening during your training. It provides tracking and visualizing metrics such as loss and accuracy, visualizing the model graph (ops and layers), viewing histograms of weights, biases, or other tensors as they change over time, displaying images, text, and audio data and much more.

3. **Weights & biases (https://www.wandb.com/)** : This is an experiment tracking tool for machine learning. Its background process collects useful data about what is happening as you train your models. For example, it can track model performance metrics, hyperparameters, gradients, system metrics, output files, and your most recent git commit. Again, like in case of the Albumentations library, Weights & biases is framework agnostic.

4. **MMDetection (https://github.com/open-mmlab/mmdetection)** : MMDetection is an object detection toolbox based on PyTorch that contains a rich set of object detection and instance segmentation methods. It started from a codebase of MMDet team who won the detection track of COCO Challenge 2018. It gradually evolves into a unified platform that covers many popular detection methods. By now it provides weights for more than 200 network models.

5. **Fast.ai (https://www.fast.ai/)** : The fast.ai library could be thought of as an independent framework, based on PyTorch. It simplifies training fast and accurate neural nets. For doing so, it uses modern best practices, including "out of the box" support for vision, text, tabular, and collaborative filtering models.

Web APIs

1. **Alchemy API :** A deep learning based API for auto tagging images based on the content of the image. If you upload an image of a cat, it will return "cat" as a tag. Deep learning based large-scale recognition is a hot topic of research these days. If you have been following **ImageNet Large Scale Visual Recognition Challenge (ILSVRC)**, you probably know that even though IBM is first to market with its API, several other teams from Google, Facebook, Microsoft, Baidu, and several universities are doing much better in the competition. Hope they come up with an API too!

2. **CloudSight** (**http://cloudsightapi.com/**) : What is better than computer vision ? Well, human vision! CloudSight API does visual recognition using a combination of computer vision and human crowd sourcing. You can use their app called CamFind to see how well it works.

3. **Face++** (**http://www.faceplusplus.com/**) : An API for face detection, facial landmark detection, face search, and face recognition.

4. **TinEye** (**https://services.tineye.com/TinEyeAPI**) : Search the entire web for an image using TinEye's reverse image search.

5. **OCRSDK** (**http://ocrsdk.com**) : Upload an image containing text and get back the results as text. They provide sample code and it works well for standard scanned text.

6. **CloudCV** (**http://cloudcv.org**) : CloudCV describes itself as a Large-Scale Distributed Computer Vision as a Cloud Service. It is not a commercial product, but is being developed by Machine Learning and Perception Lab at Virginia Tech. They do image stitching and object detection / classification in the cloud.

My Contact Info

Website	https://www.learnopencv.com
Email	spmallick@learnopencv.com
LinkedIn	https://www.linkedin.com/in/satyamallick