

# 掌握OpenCV 4- Python

## 前言

这是一本关于OpenCV计算机视觉和Python语言的书。OpenCV是一个计算机视觉库，也是一个机器学习库。为什么选OpenCV和Python?这是一个很好的问题，本书第一章中将讨论这个问题。OpenCV是最好的开源计算机视觉库(BSD许可，对学术和商业使用都是免费)，提供超过2500个优化算法，包括先进的计算机视觉算法，提供机器学习和深度学习支持。OpenCV是用优化的C/C++编写的，提供了Python封装，Python程序中可以使用。通常认为Python是科学计算的理想语言，可快速原型设计，为计算机视觉项目提供了许多预编译库。

本书将使用Matplotlib、scikit-image、SciPy、dlib、face-recognition、Pillow、cvlib、Keras、TensorFlow和Flask等库。

本书是通过Python和OpenCV应用各种技术创建高级应用程序，有较完整资源，这些技术包括面部识别、目标跟踪、增强现实、对象检测和分类等。探讨了机器学习和深度学习技术在Python生态系统有关计算机视觉的应用

读者对象主要是具有基本Python编程知识的学生，研究人员，和开发人员，从零开始学计算机视觉，并打算深入研究这个领域。假设读者有一些Python编程经验，对图像处理基本理解，如像素和颜色通道，也会有帮助，但不是必须的，这些概念在书中有所涉及。需要基本的数学知识。

## 本书内容

第1章 OpenCV设置。介绍安装Python和OpenCV编程所需的程序，OpenCV与计算机视觉的主要概念相关基础有关的专业术语和概念

- 第2章 OpenCV图像基础。如何编写第一个程序，介绍OpenCV库
- 第3章 文件和图像处理。介绍如何处理文件和图像。这是计算机视觉应用程序必需的。
- 第4章 OpenCV基本图形。介绍如何使用OpenCV库绘制图形，包括基本图形，高级形状
- 第5章 图像处理技术。介绍计算机视觉项目所需要的一般的图像处理技术。
- 第6章 直方图构建。介绍如何创建和理解直方图。这是图像内容理解的重要工具。

- 第7章 阈值技术。介绍阈值技术。计算机视觉应用中图像分割的关键过程。
- 第8章 轮廓检测、滤波和绘图。介绍轮廓处理，物体检测和识别。
- 第9章 增强现实。介绍第增强现实应用程序
- 第10章 OpenCV机器学习。进入机器学习的世界，介绍在计算机视觉项目中应用机器学习
- 第11章 人脸检测、跟踪和识别。介绍用算法进行与人脸检测、跟踪和识别相关的人脸识别项目
- 第12章 深度学习。介绍OpenCV有关的深度学习，以及TensorFlow、Keras等Python深度学习库
- 第13章 Python和OpenCV创建移动和Web计算机视觉。介绍使用Flask创建计算机视觉和深度学习Web应用程序。

## 代码下载

下载本书示例代码的网址 [www.packt.com](http://www.packt.com),

<https://github.com/PacktPublishing/Mastering-OpenCV-4-with-Python>。

## 第1部分：OpenCV4和Python简介

介绍OpenCV库，Python和OpenCV编程所需安装的库模块，第一个程序脚本，如何处理简单文件和图像。

## OpenCV设置

介绍第编程语言python，计算机视觉和机器学习库opencv，OpenCV和Python组合具有构建各种计算机应用程序的潜力。第一个Python OpenCV项目。

本章涉及：

OpenCV库理论介绍

Python OpenCV及其他程序安装

运行示例、文档、帮助的包

# 技术背景要求

Python安装，详见<https://www.python.org>。OpenCV 安装详见 <https://opencv.org>

科学计算、数据计算相关的Python包，如，NumPy ，详见 <http://www.numpy.org/>，或 [Matplotlib](https://matplotlib.org/) 详见<https://matplotlib.org/>

建议安装集成开发环境(IDE)。我们使用visual studio code

建议安装Git client。参阅[https:// education.github.com/gitchea-she-educ.pdf](https://education.github.com/gitchea-she-educ.pdf)。本书的GitHub，<https://github.com/PacktPublishing/Mastering-OpenCV-4-with-Python>。

GitHub存储库的README文件包括以下内容：代码测试规范，硬件规范，相关书籍和产品

- 第1章, OpenCV配置: opencv-contrib-python
- 第2章, OpenCV图像基础: opencv-contrib-python和 matplotlib
- 第3章, 文件和图像处理: opencv-contrib-python和 matplotlib
- 第4章, 构建OpenCV的基本形状: opencv-contrib-python 和matplotlib
- 第5章, 图像处理技术: opencv-contrib-python和 matplotlib
- 第6章, 构建和构建直方图: opencv-contrib-python 和matplotlib
- 第7章, 阈值技术: OpenCV -contrib-python, matplotlib, scikit-image, scipy
- 第8章, 轮廓检测, 滤波和绘图: OpenCV -contrib-python和matplotlib
- 第9章, 增强现实: OpenCV -contrib-python和matplotlib
- 第10章, 机器学习与OpenCV: OpenCV -contrib-python和matplotlib
- 第11章, 人脸检测, 跟踪和识别: opencv-contrib-python、matplotlib dlib, face-recognition, cvlib, requests, progressbar, keras, tensorflow
- 第12章, 介绍了深度学习: opencv-contrib-python, matplotlib, tensorflow, keras
- 第13章, OpenCV Python移动和Web计算机视觉 : opencv-contrib-python, matplotlib, flask, tensorflow, keras, requests, and pillow

运行以下命令来安装main和contrib模块

安装opencv-contrib-python

```
pip install opencv-contrib-python==4.0.0.21
```

OpenCV需要numpy。在安装opencv-contrib-python==4.0.0.21时应已安装numpy-1.16.1

- 安装Matplotlib

```
pip install matplotlib==3.0.2
```

matplotlib 需先安装kiwisolver, pyparsing, six, cycler, and python-dateutil. cycler-0.10.0, kiwisolver-1.0.1, pyparsing-2.3.1, python-dateutil-2.8.0, and six-1.12.0

- 安装图像处理算法库
- 安装 scikit-image

```
pip install scikit-image==0.14.2
```

scikit-image需要cloudpickle、decorator、networkx、numpy、toolz、dask、pillow、PyWavelets和six。

安装scikit-image==0.14.2时同时安装了PyWavelets-1.0.1, cloudpickle-0.8.0, dask-1.1.1, decorator-4.3.2, networkx-2.2, numpy-1.16.1, pillow-5.4.1, six-1.12.0, toolz-0.9.0

- 安装SciPy

```
pip install scipy==1.2.1
```

安装scipy==1.2.1时,numpy-1.16.1先安装。

- 安装dlib库

```
pip install dlib==19.8.1
```

- 安装face-recognition

```
pip install face-recognition==1.2.3
```

- Install cvlib:

```
pip install cvlib==0.1.8
pip install requests==2.21.0
pip install progressbar==2.5
pip install keras==2.2.4
pip install tensorflow==1.12.0
pip install flask==1.0.2
```

## 要求硬件性能

- 32-bit or 64-bit architecture
- 2+ GHz CPU
- 4 GB RAM
- 至少10G可用硬盘

## 了解Python

Python是解释型高级通用编程语言，具有动态类型系统和自动内存管理。Python官方主页<https://www.python.org/>。Python普及程度在过去十年中稳步上升。Python人工智能(AI)、机器学习、神经网络、深度学习、物联网(IoT)和机器人等有重要作用。

Python被认为是科学计算的语言，主要有四个原因：

- 容易理解。
- 通过包, 支持科学计算。
- 消除了其他编程语言不少复杂性。
- 有简单一致的语法。

Python刺激了快速原型开发，它有助于简化代码的编写和执行。与其他编程语言相比，Python只需五分之一的代码就可以实现相同的逻辑。Python有很多预编译库，如NumPy、SciPy、scikit-learn满足AI项目的需求。Python得益于用于科学计算的丰富生态系统，一个独立的平台，允许开发人员在不同的平台上节省测试时间。Python还提供Jupyter Notebook等工具，可简单而舒适共享脚本，有助于交互式计算环境中的协作。

# OpenCV介绍

OpenCV是C++编程库，具有实时功能。用优化C/C++编写。开源的计算机视觉库

- OpenCV (BSD许可[https://en.wikipedia.org/wiki/BSD\\_licenses](https://en.wikipedia.org/wiki/BSD_licenses))是免费图像处理库
- 有超过2500个优化算法，包括最先进的计算机视觉算法
- 机器学习和深度学习支持
- 这个库是为性能而优化的
- 有大型的开发人员社区正在使用和支持OpenCV
- 有c++、Python、Java和MATLAB接口
- 支持Windows、Linux、Android和macOS多种操作系统
- 快速和定期更新(现在每六个月发布一次正式版本)

## 背景知识

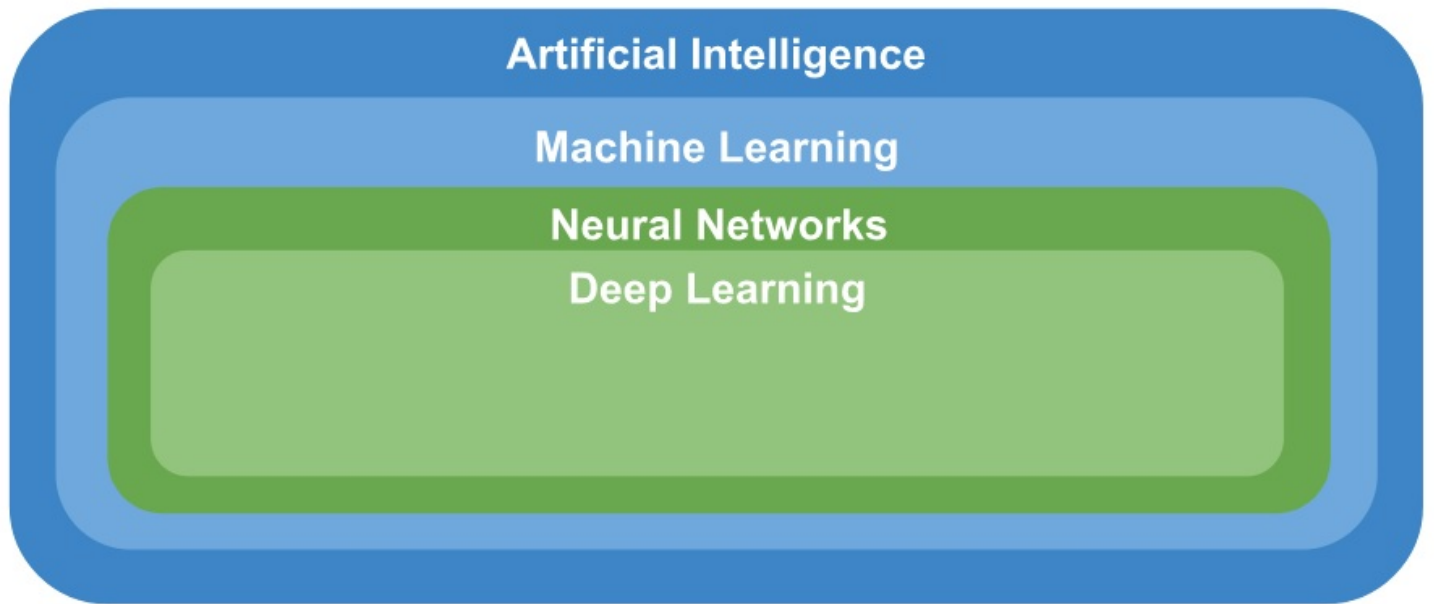
对本书主题相关的主要概念进行一些梳理。过去几年人们对人工智能和机器学习产生了浓厚的兴趣，尤其是在深度学习领域。这些说法经常互换使用，混淆。人工智能AI指的是使机器、计算机或机器人系统能够像人类一样处理信息的技术。AI常用作提供智能的机器技术的伞式结构，涵盖广泛的方法和算法。

机器学习是让计算机从历史数据中学习，并对新数据做出预测的过程。机器学习是人工智能的一个分支学科，关于机器在学习关系上使用的统计技术。在收集数据的基础上，计算机独立地学习算法。这些算法和方法包括支持向量机、决策树、随机森林、逻辑回归、贝叶斯网络和神经网络等。

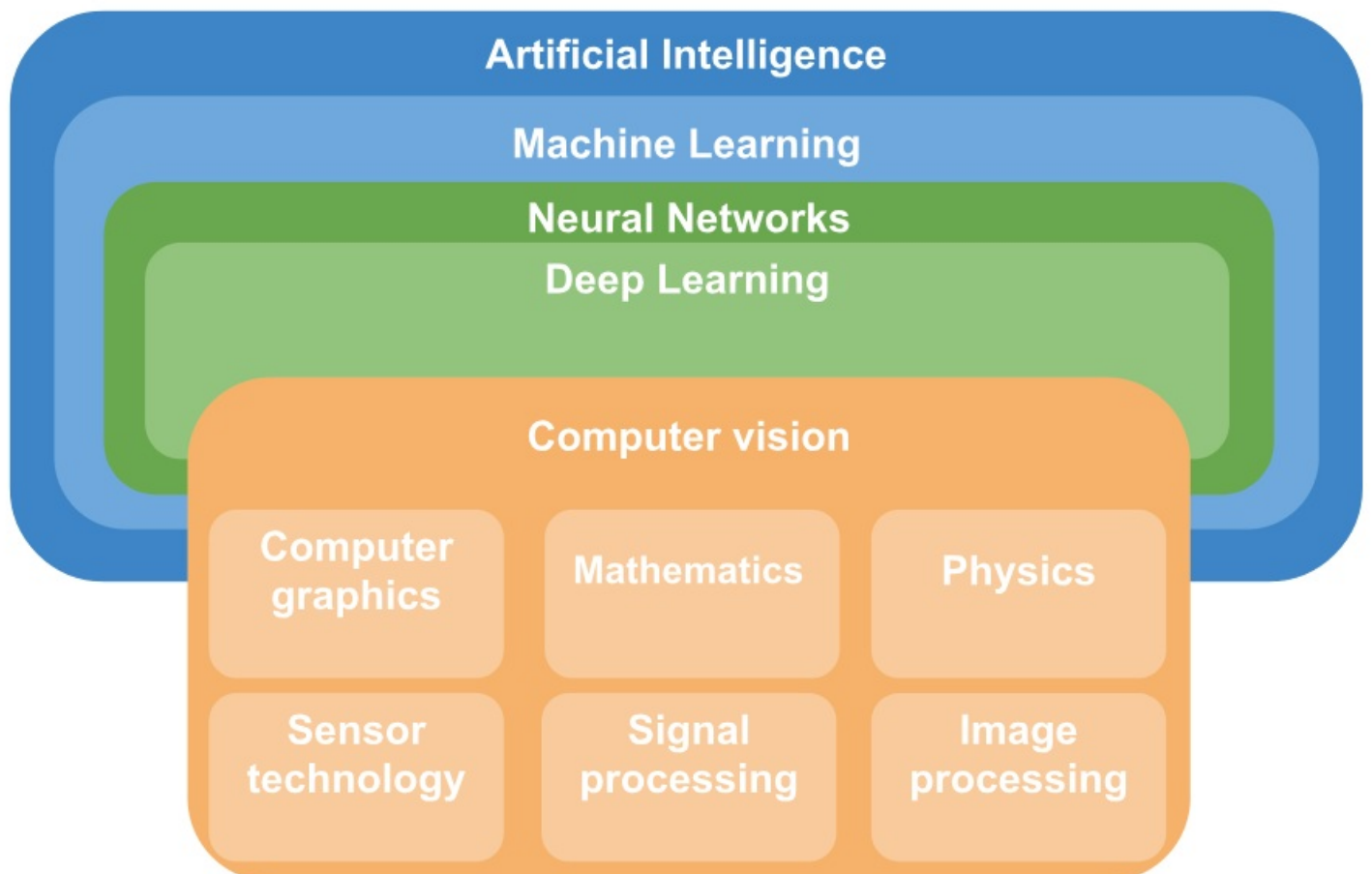
神经网络是基于生物大脑的结构和功能的机器学习的计算机模型。人工神经元处理多个输入信号，当这些输入信号的和超过一定的阈值时，向相邻的神经元发送信号。

深度学习是机器学习的一个子集，它处理大量的非结构化数据，如人类语音、文本和图像。深度学习模型是对数据进行多层数学计算的人工神经网络，其中一层的结果作为输入输入到下一层，对输入数据进行分类和/或预测。

这些概念在层次上是相互依赖的，AI是最广泛，而深度学习是最具体的，参见下图



计算机视觉是跨学科的人工智能领域，旨在为计算机和其他具有计算能力的设备，从数字图像和视频中提供高层次的理解，包括获取、处理和分析数字图像的功能依赖机器学习和深度学习算法来构建计算机视觉应用，计算机视觉在一定程度上属人工智能领域。计算机视觉由多种技术组成，包括计算机图形学、图像处理、信号处理、传感器技术、数学甚至物理学。计算机视觉学科参见下图

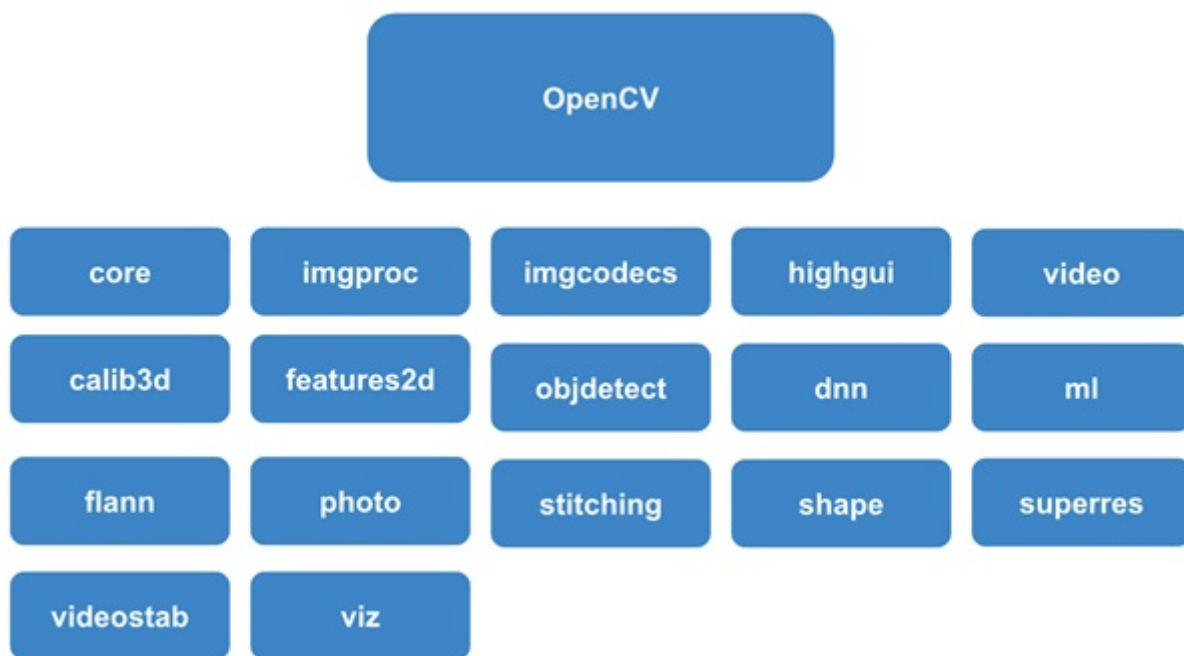


# OpenCV理论介绍

OpenCV是一个具有实时计算机视觉功能的程序库，对学术和商业使用都是免费的(BSD许可)。

## OpenCV 模块

从版本2起OpenCV分成几个模块，每个模块通常理解为专门处理一组计算机视觉问题。模块划分参见下图



### 简要描述OpenCV模块

- **core**: 核心功能。核心功能是定义基本数据结构的模块，也是库中所有其他模块使用的基本功能。
- **imgproc**: 图像处理。一种图像处理模块，包括图像滤波、几何图像变换、颜色空间变换和直方图。
- **imgcodecs**: 图像编解码器。图像文件的读写。
- **videoio**: 视频I / O。一个视频捕获和视频编解码器的接口。
- **highgui**: 高级GUI。UI功能的接口。它提供了一个接口来方便地完成以下工作
  - 创建和操作可以显示/显示图像的窗口
  - 向窗口添加跟踪条、键盘命令和处理鼠标事件



- **video**: 视频分析。视频分析模块，包括背景减除、运动估计和目标跟踪算法。
- **calib3d**: 摄像机标定和三维重建。摄像机标定和三维重建，包括基本的多视点几何算法、立体对应算法、物体姿态估计、单、立体摄像机标定和三维重建。2D功能框架。该模块包括特征检测器、描述符和描述符匹配器。
- **objdetect**: 目标检测。检测对象和预定义类的实例(例如，面孔、眼睛、人员和汽车)。
- **dnn**: 深度神经网络(dnn)模块。此模块包含以下内容
  - 用于创建新层的API
  - 一组构建的有用层
  - 从层构造和修改神经网络的API
  - 从不同的深度学习框架加载序列化网络模型的功能
- **ml**: 机器学习。机器学习库(MLL)是一组用于分类、回归和聚类的类和方法。
- **flann**: 多维空间中的聚类 and 搜索。近似近邻快速库(FLANN)是一组非常适合于快速近邻搜索的算法。
- **photo**: 计算摄影。本模块提供了计算摄影的一些功能。
- **stitching**: 图像缝合。该模块实现了一种自动全景图像拼接流水线。
- **shape**: 形状距离和匹配。形状距离和匹配模块，可用于形状匹配、检索或比较。
- **superres**: 超分辨率。此模块包含一组可用于增强分辨率的类和方法。
- **videostab**: 视频稳定。该模块包含一组用于视频稳定的类和方法。
- **viz**: 3D可视化工具。此模块用于显示小部件，这些小部件提供了几个与场景和小部件交互的方法。

## OpenCV 用户

无论专业软件开发人员还是新手，OpenCV库对于图像处理和计算机视觉领域的研究生、研究人员和计算机程序员来说都是非常有趣的。该库提供了许多最先进的计算机视觉算法，受到科学家和学者的欢迎。

OpenCV常用作计算机视觉和机器学习的教学工具。OpenCV足够健壮，可以支持应用程序开发，广泛用于非商业和商业产品。如谷歌、微软、英特尔、IBM、索尼和本田等公司都在使用OpenCV。一流大学的研究机构，如麻省理工学院、卡内基梅隆大学或斯坦福大学，为库提供支持。OpenCV有超过1400万的下载量和超过47000的用户。

## OpenCV 应用，如

- OpenCV的应用范围非常广泛
- 2 D和3 D功能工具箱
- 街景图像拼接
- Egomotion估计
- Facial-recognition估计
- 人脸识别系统
- 人机交互手势识别
- 移动机器人
- 运动理解
- 目标识别
- 自动检查和监测
- 分割和识别
- 立体视觉深度感知双摄像头
- 医学图像分析
- 运动结构
- 运动跟踪
- 增强现实
- 视频/图片搜索和检索
- 导航和机器人和无人驾驶汽车控制
- 驾驶员睡意和注意力分散

## 为什么在研究中引用OpenCV

如果在研究中使用OpenCV，建议引用OpenCV库。这样，其他研究人员可以更好地理解你提出的算法，并重现你的结果，以获得更好的可信度。OpenCV也增加反响，从而产生更好的计算机视觉库。下面的代码显示了引用OpenCV的BibTex条目

```
@article{opencv_library,  
author = {Bradski, G.},  
citeulike-article-id = {2236121},  
journal = {Dr. Dobb's Journal of Software Tools},  
keywords = {bibtex-import},  
posted-at = {2008-01-15 19:21:54},  
priority = {4},
```

```
title = [{The OpenCV Library}],
year = {2000}
}
```

# OpenCV包的主要用途

图像处理: NumPy, OpenCV, scikit-image, PIL Pillow, SimpleCV, Mahotas, ilastik

文本处理: NLTK, spaCy, NumPy, scikit-learn, PyTorch

音频处理: LibROSA

机器学习问题: pandas, scikit-learn, Orange, PyBrain, Milk

数据可视化: Matplotlib, Seaborn, scikit-learn, Orange

深度学习: TensorFlow, Pytorch, Theano, Keras

- NumPy 提供对大型多维数组的支持。NumPy是计算机视觉中的一个关键库，因为图像可以表示为多维数组。将图像表示为NumPy数组有很多优点。<http://www.numpy.org/>)
- OpenCV是开源的计算机视觉库。<http://www.numpy.org/>)
- Scikit-image图像处理算法集。scikit-image图像处理由简单的Numpy数组完成。  
<https://scikit-image.org/>
- Python图像库(PIL) 是图像处理库，提供强大的图像处理和图形功能。  
<http://www.pythonware.com/products/pil/>
- Pillow 给Python解释器增加图像处理能力。<https://pillow.readthedocs.io>
- SimpleCV是提供图像处理关键功能的计算机视觉框架。<http://simplecv.org>
- Mahotas是Python图像处理和计算机视觉函数，最初是生物图像信息学设计，在其他领域也有用途。数据类型完全基于numpy数组。<https://mahotas.readthedocs.io>
- Ilastik是用于交互式图像分割、分类和分析的简单工具。<http://ilastik.org/>
- Scikit-learn以各种分类、回归和聚类算法为特色的机器学习库。 <http://scikit-learn.org>
- SciPy 是科学和技术计算库。<https://www.scipy.org/>
- NLTK是处理人类语言数据的库和程序。 <https://www.nltk.org/>
- spaCy是一个用于Python自然语言处理的开源软件库。 <https://spacy.io/>
- LibROSA是一个音乐和音频处理库。 <https://librosa.github.io/librosa/>
- Panda 构建在NumPy之上，提供高级数据计算工具和数据结构的库。  
<https://pandas.pydata.org>
- Matplotlib 生成各种格式图形的绘图库。<https://matplotlib.org/>

- Seaborn 构建在Matplotlib之上的图形库。(<https://seaborn.pydata.org/>)
- Orange 面向新手和专家的开源机器学习和数据可视化工具包。  
(<https://orange.biolab.si/>)
- PyBrain 为机器学习提供易于使用的最新算法的机器学习库。(<http://pybrain.org/>)
- Milk是机器学习工具包，主要关注有多个分类器的监督分类。  
(<http://luispedro.org/software/milk/>)
- TensorFlow 开源的机器学习和深度学习库。(<https://www.tensorflow.org/>)
- PyTorch 开源的机器学习和深度学习库。(<https://pytorch.org/>)
- Theano 用于快速数学表达式、计算和计算的库，已被编译在CPU和GPU架构上运行。  
(<http://deeplearning.net/software/theano/>)
- Keras 高级深度学习库，可以运行在TensorFlow、CNTK、Theano或Microsoft Cognitive Toolkit之上。(<https://keras.io/>)
- Django 基于python的开源web框架，快速开发和简洁、实用的设计。  
(<https://www.djangoproject.com/>)
- Flask 用Python编写的基于Werkzeug和Jinja 2的web框架。  
(<http://flask.pocoo.org/>)

\*更多的人工智能和机器学习的Python库和包，可参看

<https://python.libhunt.com/packages/artificial-intelligence>

## OpenCV和Python项目结构

安装测试

```
python test_opencv_installation.py
```