目录

一、hjimce 算法博文目录

1、**深度学习**

深度学习（一）深度学习学习资料

深度学习（二）theano 环境搭建

深度学习（三）theano 入门学习

深度学习（四）卷积神经网络入门学习

深度学习（五）caffe 环境搭建

深度学习（六）caffe 入门学习

深度学习（七）caffe 源码 c++学习笔记

深度学习（八）RBM 受限波尔兹曼机学习-未完待续

深度学习（九）caffe 预测、特征可视化 Python 接口调用

深度学习（十）keras 学习笔记

深度学习（十一）RNN 入门学习

深度学习（十二）从自编码到栈式自编码

深度学习（十三）caffe 之训练数据格式

深度学习（十四）基于 CNN 的性别、年龄识别

深度学习（十五）基于级联卷积神经网络的人脸特征点定位-CVPR 2013

深度学习（十六）基于 2-channel network 的图片相似度判别-CVPR 2015

深度学习（十七）基于改进 Coarse-to-fine CNN 网络的人脸特征点定位-ICCV 2013

深度学习（十八）基于 R-CNN 的物体检测-CVPR 2014

深度学习（十九）基于空间金字塔池化的卷积神经网络物体检测-ECCV 2014

深度学习（二十）基于 Overfeat 的图片分类、定位、检测-2014 ICLR

深度学习（四十二）word2vec 词向量学习笔记

2、机器学习基础实践

机器学习（一）线性回归

机器学习（二）逻辑回归

机器学习（三）k 均值聚类

机器学习（四）高斯混合模型

机器学习（五）PCA 数据降维

机器学习（六）非负矩阵分解 NMF-未完待续

机器学习（七）白化 whitening

机器学习（八）Apriori 算法学习

机器学习（九）初识 BP 神经网络

机器学习（十）Mean Shift 聚类算法

机器学习（十一）谱聚类算法

机器学习（十二）朴素贝叶斯分类

机器学习（十三）k-svd 字典学习

机器学习（十四）Libsvm 学习笔记

3、图像处理

图像处理（一）图像变形(1)矩形全景图像还原-Siggraph 2014

图像处理（二）Seam Carving 算法-Siggraph 2007

图像处理（三）图像分割(1)Random Walks 分割

图像处理（四）图像分割(2)测地距离 Geodesic 图割

图像处理（六）递归双边滤波磨皮

图像处理（七）导向滤波磨皮

4、三维图形

5、基础知识

基础知识（一）matlab 与 c++混合编程之环境搭建

基础知识（二）matlab 与 c++混合编程之经验笔记

基础知识（三）makefile 文件编写初级篇

基础知识（四）Dijkstra 算法

基础知识（五）对齐变换相关函数

基础知识（六）KD-Tree 快速最近邻搜索

基础知识（七）opencv、python、人脸框检测

基础知识（八）python 文件操作相关函数

基础知识（九）boost+vs2015 安装配置

基础知识（十）C++常用函数.txt

基础知识（十一）Eigen 求解稀疏矩阵

基础知识（十二）Opengl 回顾记录

基础知识(十三)dlib python 人脸检测 特征点定位

基础知识(十四)服务器搭建

基础知识（十五）Cmake

基础知识（十六）Opencv、python、ubuntu

二、机器视觉开源代码集合

1、特征提取 Feature Extraction

- SIFT[1] [Demo program][SIFT Library] [VLFeat]

- PCA-SIFT [2] [Project]

- Affine-SIFT[3] [Project]

- SURF[4] [OpenSURF] [Matlab Wrapper]

- Affine Covariant Features[5] [Oxford project]

- MSER [6] [Oxford project] [VLFeat]

- Geometric Blur [7] [Code]

- Local Self-Similarity Descriptor [Oxford implementation]

- Global and Efficient Self-Similarity[9] [Code]

- Histogram of Oriented Graidents [10] [INRIA Object Localization Toolkit] [OLT toolkit for Windows]

- GIST [11] [Project]

- Shape Context [12] [Project]

- Color Descriptor [13] [Project]

- Pyramids of Histogram of Oriented Gradients[Code]

- Space-Time Interest Points(STIP) [14][Project] [Code]

- Boundary Preserving Dense Local Regions [15][Project]

- Weighted Histogram[Code]

- Histogram-based Interest Points Detector[Paper][Code]

- An OpenCV C++ implementation of Local Self Similarity Descriptors [Project]

- Fast Sparse Rpresentation with PrototypesProject]

- Corner Detection[Project]

- AGAST Corner Detector:faster than FAST and even FAST-ER[Project]

- Global and Efficient Self-Similarity for Object Classification and Detection[code]

- WαSH: Weighted α-Shape for Local Feature Detection[Project]

- HOG[Project]

- Online Selection of Discriminative Tracking Features[Project]

2、图像分割 Image Segmentation

- Normalized Cut[1] [Matlab code]

- Gerg Mori' Superpixel code [2] [Matlab code]

- Efficient Graph-based Image Segmentation   [3] [C++ code] [Matlab wrapper]

- Mean-Shift Image Segmentation [4] [EDISON C++ code] [Matlab wrapper]

- OWT-UCM Hierarchial Segmentation [5] [Resources]

- Turbepixels [6] [Matlab code 32bit] [Matlab code 64bit] [Updated code]

- Quick-Shift [7] [VLFeat]

- SLIC Superpixels [8] [Project]

- Segmentation by Minimum Code Length [9] [Project]

- Biased Normalized Cut [10] [Project]

- Segmentation Tree [11-12] [Project]

- Entory Rate Superpixel Segmentation[13] [Code]

- Fat Approximate Energy Minimization via Graph Cuts[Paper][Code]

- Efficient Planer Graph Cuts with Applications in Computer Vision[Paper][Code]

- Isoperimetric Graph Partitioning for Image Segmentation[Paper][Code]

- Random Walks for Image Segmentation[Paper][Code]

- Blossom V: A new implementation of minimum cost perfect matching algorithm [Code]

- An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Computer Vision[Paper][Code]

- Geodesic Star Convexity for Interactive Image Segmentation [Project]

- Biased Noralized Cuts[Project]

- Max-flow/min-cut[Project]

- Chan-Vese Segmentation using Level Set[Project]

- A Toolbox of Level Set Methods[Project]

- Re-intialization Free Level Set Evolution via Reaction Diffusion[Project]

- Improved C-V active contour mode[Paper][Code]

- A variational Multiphase Level Set Approach to Simultaneous Segmentation and Bias Correction[Paper][Code]

- Level Set Method Research by Chunming Li[Project]

- ClassCut for Unsupervised Class Segmentation[code]

- SEEDS: Superpixels Extracted via Energy-Driven Sampling[Project][other]

3、目标检测 Object Detection

- A simpled object detector with boosting [Project]

- INRIA Object Detection and Localization Toolkit[1] [Project]

- Discriminatively Trained Deformable Part Models [2] [Project]

- Cascase Object Detection with Deformable Part Models [3] [Project]

- Poselet[4] [Project]

- Implicit Shape Model[5] [Project]

- Viola and Jones's Face Detection [6] [Project]

- Bayesian Modelling of Dyanmic Scenes for Object Detection[Paper][Code]

- Hand detection using multiple proposals[Project]

- Color Constancy,Intrinsic Image,and Shape Estimation[Paper][Code]

- Discriminatively trained deformable part models[Project]

- Gradient Reponse Maps for Real-Time Detection of Texture-Less Objects:LineMOD [Project]

- Image Processing On Line[Project]

- Robust Optical Flow Estimation[Project]

- Where's Waldo:Matching People in Images of Crowds[Project]

- Scalable Multi-class Object Detection[Project]

- Class-Specific Hough Forests for Object Detection[Project]

- Deformed Lattice Detection In Real-World Imaged[Project]

- Discriminatively trained deformable part models[Project]

4、显著性检测 Saliency Detection

- Itti,Koch,and Niebur' saliency detection[1] [Matlab code]

- Frequency-tuned salient region detection[2] [Project]

- Saliency detection using maximum symmetric surround [3] [Project]

- Attention via Information Maximization [4] [Matlab code]

- Context-aware saliency detection [5] [Matlab code]

- Graph-based visual saliency[6] [Matlab code]

- Saliency detection: A speclral residual approach[7] [Matlab code]

- Segmenting salient objects from images and videos[8] [Matlab code]

- Saliency Using Natural statistics [9] [Matlab code]

- Discriminant Saliency for Visua Recognition from Cluttered Scenes. [10] [Code]

- Learning to Predict Where Humans Look[11] [Project]

- Global Contrast based Salient Salient Region Detection [12] [Project]

- Bayesian Saliency via Low and Mid Level Cues[Project]

- Top-Down Visual Saliency via Joint CRF and Dictionary Learning [Paper][Code]

- Saliency Detection: A Spectral Residual Approach

5、图像分类、聚类

- Pyramid Match[1] [Project]

- Spatial Pyramid Matching   [2] [Code]

- Locality-constrained Linear Coding[3] [Project] [Matlab code]

- Sparse Coding   [4] [Project] [Matlab code]

- Texture Classification[5] [Project]

- Multiple Kernels for Image Classification [6] [Project]

- Feature Combination[7] [Project]

- SuperParsing   [Code]

- Large Scale Correlating Clustering Optimization[Matlab code]

- Detection and Sketching the Commom[Project]

- Self-Tuning Spectral Clustering[Project][Code]

- User Assisted Separation of Reflections from a Single Image Using a Sparsity Prior [Paper][Code]

- Filters for Texture Classfication[Project]

- Multiple Kernel Learning for Image Classification[Project]

- SLIC Superpixels[Project]

6、抠图 Image Matting

- A Closed Form Solution to Natural Image Matting[Code]

- Spectral Matting [Project]

- Learning-based Matting [Code]

7、目标跟踪 Object Tracking

- A Forest of Sensors-Tracking Adaptive Background Mixture Models[Project]

- Object Tracking via Partial Least Squares Analysis[Paper][Code]

- Robust Object Tracking with Online Multiple Instance Learing[Paper][Code]

- Online Visual Tracking with Histograms and Articulating Blocks[Project]

- Incremental Learning for Robust Visual Tracking[Project]

- Real-time Compressive Tracking[Project]

- Robust Object Tracking via Sparsity-baased Collaborative Model[Project]

- Visual Tracking via Adapative Structural Local Sparse Appearance Model [Project]

- Online Discriminative Object Trackign with Local Sparse Representation[Paper][Code]

- Superpixel Tracking[Project]

- Learning Hierarchial Image Represetation with Sparsity, Saliency and Locality[Paper][Code]

- Online Multiple Support Instance Tracking [Paper][Code]

- Visual Tracking with Online Multiple Instance Learning[Project]

- Object detection and recognition[Project]

- Compressive Sensing Resources[Project]

- Robust Real-time Visual Tracking using Pixel-Wise Posteriors[Project]

- Tracking-Learning-Detection[Project] [OpenTLD/C++ Code]

- The HandVu: vision-based hand gesture interface[Project]

- Learning Probabilistic Non-Linear Latent Variable Models for Tracking Complex Activities[Project]

8、Kinect

- Kinect toolbox [Project]

- OpenNI[Project]

- Zouxy09 CSDN Blog [Resource]

- Finger Tracker 手指跟踪[code]

9、3D 相关

- 3D Reconstruction of a Moving Object[Paper] [Code]

- Shape From Shading Using Linear Approximation[Code]

- Combining Shape from Shading and Stereo Depth Maps[Project][Code]

- Shape from Shading :A Survey [Paper][Code]

- A Spatio-Temporal Descriptor based on 3D Gradients (HOG3D) [Project][Code]

- Multi-camera Scene Reconstruction via Graph Cuts [Paper][Code]

- A Fast Marching Formulation of Perspective Shape from Shading under Frontal Illumination [Paper][Code]

- Reconstruction:3D Shape ,Illumination, Shading, Reflectance,Texture [Project]

- Monocular Tracking of 3D Human Motion with a Coordianted Mixtyre of Factor Analyzers[Code]

- Learning 3-D Scene Structure from a Single Still Image[Project]

10、机器学习算法

- Matlab class for computing Approximate Nearest Nieghbor (ANN) [Matlab class providing interface to ANN library]

- Random Sampling[code]

- Probabilistic Latent Semantic Analysis(pLSA)[Code]

- FASTANN and FASTCLUSTER for approximate k-means (AKM) [Project]

- Fast Intersection /Additive Kernel SVMS[Project]

- SVM[Code]

- Ensemble Learning[Project]

- Deep Learning[Net]

- Deep Lerning Methods for Vision[Project]

- Neural Network for Recognition of Handwritten Digits[Project]

- Training a deep autoencoder or a classifier on MNIST digits[Project]

- The MNIST database of handwritten digits[Project]

- Ersatz: deep neural networks in the cloud[Project]

- Deep Learning [Project]

- sparseLM: Sparse Levenberg-Marquardt nonlinear least squares in C/C++[Project]

- Weka 3:Deta Mining Software in java[Project]

- Invited talk "A tutorial on Deep Learning" by Dr.Kai Yu(余凯) [Video]

- CNN-Convolutional neural network class[Matlab Tool]

- Yann LeCun's Publications[Wedsite]

- LeNet-5,convolutional neural networks[Project]

- Training a deep autoencoder or a classifier on MINIST digits[Project]

- Deep Learning 大牛 Geoffrey E. Hinton's HomePage[Website]

- Multiple Instance Logistic Discriminant-based Metric Learning (MildML) and Logistic Discriminant-based Metric Learning(LDML) [Code]

- Sparse coding simulation software[Project]

- Visual Recognition and Machine Learning Summer School[Project]

11、目标、行为识别

- Action Recognition by Dense Trajectories[Project][Code]

- Action Recognition Using a Distributed Representation of Pose and Appearance[Project]

- Recognition Using Region[Paper][Code]

- 2D Articulated Human Pose Estimation[[Project]]

- Fast Human Pose Estimation Using Appearance and Motion via Multi-Dimensional Boosting Regression[[Paper]][[Code]]

- Estimation Human Pose from Occluded Images[[Paper]][[Code]]

- Quasi-dense wide baseline matching[[Project]]

- ChaLearn Gesture Challenge: Principal motion:PCA-based reconstruction of motion histogram[[Project]]

- Real Time Head Pose Estimation with Random Regression Forests[[Project]]

- 2D Action Recognition Serves 3D Human Pose Estimation[[Project]]

- A hough Transform-Based Voting Framework for Action Recognition[[Project]]

- Motion Interchange Patterns for Action Recognition in Unconstrained Videos[[Project]]

- 2D articulated human pose estimation software[[Project]]

- Learning and detecting shape models[[code]]

- Progressive Seaerch Space Reduction for human Pose Estimation[[Project]]

- Learning Non-Rigid 3D Shape from 2D Motion[[Project]]

12、图像处理

- Distance Transforms of Sampled Functions[[Project]]

- The Computer Vision Homepage[[Project]]

- Efficient appearance distance between windows [[code]]

- Image Exploration algorithm[[code]]

- Motion Magnification 运动放大[[code]]

- Bilateral Filtering for Gray and Color Images 双边滤波器 [Project]

- A Fast Approximation of Bilateral Filter using a Signal Processing Approach [Project]

13、人手及指尖检测与识别

- Finger-detection-and-gesture-recognition [Code]

- Hand and Finger Detection using JavaCV[Project]

- Hand and fingers detection[Code]

14、场景解释

- Nonparametric Scene Parsing Via Label Transfer [Project]

15、光流 Optical Flow

- High accuracy optical using a theory for warping [Project]

- Dense Trajectories Video Description[Project]

- SIFT Flow: Dense Correspondense across Scenes and its Applications [Project]

- KLT: An Implementation of the Kanade-Lucas-Tomasi Feature Tracker [Project]

- Tracking Cars Using Optical Flow[Project]

- Secrets of optical flow estimation and their principles[Project]

- Implmentaiton of the Black and Anandan dense optical flow method [Project]

- Optical Flow Computation[Project]

- Beyond Pixels: Exploring New Representations and Applications for Motion Analysis[Project]

- A Database and Evaluation Methodology fr Optical Flow[Project]

- Optical flow relative[Project]

- Robust Optical Flow Estimation   [Project]

- Optical flow[Project]

16、图像检索

- Semi-Supervised Distance Metric Learning for Collaborative Image Retrieval [Paper][code]

17、马尔科夫随机场

- Markov Random Fields for Super-Resolution [Project]

- A Comparative Study of Energy Minimization Methods for Markov Random Fields with Smoothness-Based Priors [Project]

18、运动检测

- Moving Object Extraction, using Models or Analysis of Regions[Project]

- Background Subtraction : Experiments and Improvements for ViBe [Project]

- A Self-Organized Approach to Background subtraction for Visual Surveillance Applications [Project]

- Changedetection.net: A new change detection benchmark dataset[Project]

- ViBe- a powerful technique for background detection and subtraction in video sequences [Project]

- Background Subtraction Program[Project]

- Motion Detection Algorithms[Project]

- Stuttgart Artificial Background Subtraction Dataset[Project]

- Object Detection, Motion Estimation, and Tracking[Project]

19、Utils 实用工具

三、开源代码文献

1、Deep Residual Networks

2、Object detection

- **Object detection via a multi-region & semantic segmentation-aware CNN model**
  https://github.com/gidariss/mrcnn-object-detection

- DeepBox: Learning Objectness with Convolutional Networks ICCV2015 proposal re-ranker
  https://github.com/weichengkuo/DeepBox

- Object-Proposal Evaluation Protocol is 'Gameable' 好多 Proposal 代码

  https://github.com/batra-mlp-lab/object-proposals

- Fast R-CNN

  https://github.com/rbgirshick/fast-rcnn

- Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks

  https://github.com/ShaoqingRen/faster_rcnn MATLAB

  https://github.com/rbgirshick/py-faster-rcnn **Python**

- YOLO : Real-Time Object Detection

  http://pjreddie.com/darknet/yolo/

  https://github.com/pjreddie/darknet

- SSD: Single Shot MultiBox Detector 比 Faster R-CNN 又快又好啊！

  https://github.com/weiliu89/caffe/tree/ssd

- A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection

  https://github.com/zhaoweicai/mscnn

3、Image Question Answering

- Stacked Attention Networks for Image Question Answering CVPR2016

  https://github.com/zcyang/imageqa-san

- Image Question Answering using Convolutional Neural Networ with Dynamic Parameter Prediction CVPR2016

- 项目网页

  http://cvlab.postech.ac.kr/research/dppnet/

  开源代码

  https://github.com/HyeonwooNoh/DPPnet

- **场景识别：

  SegNet: A Deep Convolutional Encoder-Decoder Architecture for Robust

Semantic Pixel-Wise Labelling

https://github.com/alexgkendall/caffe-segnet

4、Tracking

- Learning to Track: Online Multi-Object Tracking by Decision Making ICCV2015
  使用 Markov Decision Processes 做跟踪，速度可能比较慢,效果应该还可以
  https://github.com/yuxng/MDP_Tracking

- Fully-Convolutional Siamese Networks for Object Tracking
  http://www.robots.ox.ac.uk/~luca/siamese-fc.html

5、Car detection

- Integrating Context and Occlusion for Car Detection by Hierarchical And-or Model ECCV2014
  http://www.stat.ucla.edu/~boli/projects/context_occlusion/context_occlusion.html

6、Face detection

- 人脸检测 2015 进展：http://www.cvrobot.net/latest-progress-in-face-detection-2015/

- Face detection without bells and whistles
  project：http://markusmathias.bitbucket.org/2014_eccv_face_detection/
  Code：https://bitbucket.org/rodrigob/doppia
  Talk：http://videolectures.net/eccv2014_mathias_face_detection/ （不错的报告）

- From Facial Parts Responses to Face Detection: A Deep Learning Approach
  ICCV2015 email to get code and model
  http://personal.ie.cuhk.edu.hk/~ys014/projects/Faceness/Faceness.html

- A Fast and Accurate Unconstrained Face Detector 2015 PAMI
  简单 快速 有效
  http://www.cbsr.ia.ac.cn/users/scliao/projects/npdface/

7、Face Alignment

- Face Alignment by Coarse-to-Fine Shape Searching
  http://mmlab.ie.cuhk.edu.hk/projects/CFSS.html

- High-Fidelity Pose and Expression Normalization for Face Recognition in the Wild
  http://www.cbsr.ia.ac.cn/users/xiangyuzhu/projects/HPEN/main.htm

8、Face Recognition

- Deep face recognition
  http://www.robots.ox.ac.uk/~vgg/software/vgg_face/

- Do We Really Need to Collect Millions of Faces for Effective Face Recognition?
  http://www.openu.ac.il/home/hassner/projects/augmented_faces/

9、Person Re-identification

- Person Re-identification Results
  http://www.ssig.dcc.ufmg.br/reid-results/#ref35VIPER

- Learning a Discriminative Null Space for Person Re-identification
  code http://www.eecs.qmul.ac.uk/~lz/

- Query-Adaptive Late Fusion for Image Search and Person Re-identification CVPR2015
  http://www.liangzheng.com.cn/Project/project_fusion.html

- Efficient Person Re-identification by Hybrid Spatiogram and Covariance Descriptor CVPR2015 Workshops
  https://github.com/Myles-ZMY/HSCD

- Person Re-Identification by Iterative Re-Weighted Sparse Ranking PAMI 2015
  http://www.micc.unifi.it/masi/code/isr-re-id/ 没有特征提取代码

- Person re-identification by Local Maximal Occurrence representation and metric learning CVPR2015
  http://www.cbsr.ia.ac.cn/users/scliao/projects/lomo_xqda/

10、Head detection

- Context-aware CNNs for person head detection

  Matlab code & dataset avaiable

  http://www.di.ens.fr/willow/research/headdetection/

11、Pedestrian detection

- Pedestrian Detection with Spatially Pooled Features and Structured Ensemble

  Learning PAMI 2015

  Strengthening the Effectiveness of Pedestrian Detection with Spatially Pooled

  Features ECCV2014

  https://github.com/chhshen/pedestrian-detection

- Is Faster R-CNN Doing Well for Pedestrian Detection

  Matlab 代码 ：https://github.com/zhangliliang/RPN_BF/tree/RPN-pedestrian

12、Deep Learning

- Deeply Learned Attributes for Crowded Scene Understanding

  https://github.com/amandajshao/www_deep_crowd

  http://www.ee.cuhk.edu.hk/~jshao/WWWCrowdDataset.html

- Quantized Convolutional Neural Networks for Mobile Devices

  https://github.com/jiaxiang-wu/quantized-cnn

13、Human Pose Estimation

- DeepPose: Human Pose Estimation via Deep Neural Networks，CVPR2014

  https://github.com/mitmul/deeppose not official implementation

- Articulated Pose Estimation by a Graphical Model with Image Dependent Pairwise

  Relations NIPS 2014

  http://www.stat.ucla.edu/~xianjie.chen/projects/pose_estimation/pose_estimation.html

- Learning Human Pose Estimation Features with Convolutional Networks

  https://github.com/stencilman/deep_nets_iclr04

- Flowing ConvNets for Human Pose Estimation in Videos

  http://www.robots.ox.ac.uk/~vgg/software/cnn_heatmap/

14、杂项

- Unsupervised Learning of Visual Representations using Videos 很有前途啊！

  https://github.com/xiaolonw/caffe-video_triplet

- Learning Deep Representations of Fine-Grained Visual Descriptions

  https://github.com/reedscot/cvpr2016

- Fast Detection of Curved Edges at Low SNR

  http://www.wisdom.weizmann.ac.il/~yehonato/projectPage.html

- Unsupervised Processing of Vehicle Appearance for Automatic Understanding in Traffic Surveillance

- code: https://medusa.fit.vutbr.cz/traffic/research-topics/fine-grained-vehicle-recognition/unsupervised-processing-of-vehicle-appearance-for-automatic-understanding-in-traffic-surveillance/

15、Image Retrieval

- Learning Compact Binary Descriptors with Unsupervised Deep Neural Networks

  https://github.com/kevinlin311tw/cvpr16-deepbit

- Deep Supervised Hashing for Fast Image Retrieval

  https://github.com/lhmRyan/deep-supervised-hashing-DSH

- Bit-Scalable Deep Hashing with Regularized Similarity Learning for Image Retrieval and Person Re-identification

  https://github.com/ruixuejianfei/BitScalableDeepHash

16、数据库

- MPII Human Pose Dataset

  http://human-pose.mpi-inf.mpg.de/#overview

- WIDER FACE: A Face Detection Benchmark 数据库

http://mmlab.ie.cuhk.edu.hk/projects/WIDERFace/

17、文字识别

● FASText: Efficient Unconstrained Scene Text Detector

https://github.com/MichalBusta/FASText

18、开源车牌识别代码

● 支持美国和欧洲车牌

http://www.openalpr.com

19、DPM

● 将 voc-release4.0.1 **Linux** 转到 windows

http://blog.csdn.net/masibuaa/article/details/17577195

四、基于 DL 的计算机视觉

1、Python 基础介绍

2、实现图像分类最简单的方法：KNN

3、SVM 和 Softmax 分类器

4、SGD

5、理解反向传播

6、神经网络

7、数据的预处理、正则化

8、神经网络是如何训练的

9、神经网络之动手实践

10、详谈 CNN

11、基于 DL 的快速图像检索系统

五、CV Online 数据集

1、Action Databases

- 3D online action dataset - There are seven action categories (Microsoft and Nanyang Technological University)
- 50 Salads - fully annotated 4.5 hour dataset of RGB-D video + accelerometer data, capturing 25 people preparing two mixed salads each (Dundee University, Sebastian Stein)
- ASLAN Action similarity labeling challenge database (Orit Kliper-Gross)
- Action Detection in Videos - MERL Shopping Dataset consists of 106 videos, each of which is a sequence about 2 minutes long (Michael Jones, Tim Marks)

- [Actor and Action Dataset](#) - 3782 videos, seven classes of actors performing eight different actions (Xu, Hsieh, Xiong, Corso)
- [An analyzed collation of various labeled video datasets for action recognition](#) (Kevin Murphy)
- [BEHAVE Interacting Person Video Data with markup](#) (Scott Blunsden, Bob Fisher, Aroosha Laghaee)
- [BU-action Datasets](#) - Three image action datasets (BU101, BU101-unfiltered, BU203-unfiltered) that have 1:1 correspondence with classes of the video datasets UCF101 and ActivityNet. (S. Ma, S. A. Bargal, J. Zhang, L. Sigal, S. Sclaroff.)
- [Berkeley MHAD: A Comprehensive Multimodal Human Action Database](#) (Ferda Ofli)
- [Benchmarking Activity Recognition](#) – CVPR 2012 tutorial covering various datasets for action recognition.
- [Berkeley Multimodal Human Action Database](#) - five different modalities to expand the fields of application (University of California at Berkeley and Johns Hopkins University)
- [Breakfast dataset](#) - It's a dataset with 1712 video clips showing 10 kitchen activities, which are hand segmented into 48 atomic action classes . (H. Kuehne, A. B. Arslan and T. Serre)
- [Brown Breakfast Actions Dataset](#) - 70 hours, 4 million frames of 10 different breakfast preparation activities (Kuehne, Arslan and Serre)
- [CAD-120 dataset](#) - focuses on high level activities and object interactions (Cornell University)
- [CAD-60 dataset](#) - The CAD-60 and CAD-120 data sets comprise of RGB-D video sequences of humans performing activities (Cornell University)
- [CVBASE06: annotated sports videos](#) (Janez Pers)
- [Charades Dataset](#) - 10,000 videos from 267 volunteers, each annotated with multiple activities, captions, objects, and temporal localizations. (Sigurdsson, Varol, Wang, Laptev, Farhadi, Gupta)
- [Composable activities dataset](#) - Different combinations of 26 atomic actions formed 16 activity classes which were performed by 14 subjects and annotations were provided (Pontificia Universidad Catolica de Chile and Universidad del Norte)
- [Cornell Activity Datasets CAD 60, CAD 120](#) (Cornell Robot Learning Lab)

- DMLSmartActions dataset – Sixteen subjects performed 12 different actions in a natural manner (University of British Columbia)
- Depth-included Human Action video dataset – It contains 23 different actions (CITI in Academia Sinica)
- DogCentric Activity Dataset – first-person videos taken from a camera mounted on top of a *dog* (Michael Ryoo)
- ETS Hockey Game Event Data Set – This data set contains footage of two hockey games captured using fixed cameras. (M.-A. Carbonneau, A. J. Raymond, E. Granger, and G. Gagnon)
- FCVID: Fudan-Columbia Video Dataset – 91,223 Web videos annotated manually according to 239 categories (Jiang, Wu, Wang, Xue, Chang)
- G3D – synchronised video, depth and skeleton data for 20 gaming actions captured with Microsoft Kinect (Victoria Bloom)
- G3Di – This dataset contains 12 subjects split into 6 pairs (Kingston University)
- Gaming 3D dataset – real-time action recognition in gaming scenario (Kingston University)
- Georgia Tech Egocentric Activities - Gaze(+) – videos of where people look at and their gaze location (Fathi, Li, Rehg)
- HMDB: A Large Human Motion Database (Serre Lab)
- Hollywood 3D dataset – 650 3D video clips, across 14 action classes (Hadfield and Bowden)
- Human Actions and Scenes Dataset (Marcin Marszalek, Ivan Laptev, Cordelia Schmid)
- HumanEva: Synchronized Video and Motion Capture Dataset for Evaluation of Articulated Human Motion (Brown University)
- I-LIDS video event image dataset (Imagery library for intelligent detection systems) (Paul Hosner)
- I3DPost Multi-View Human Action Datasets (Hansung Kim)
- IAS-lab Action dataset – contain sufficient variety of actions and number of people performing the actions (IAS Lab at the University of Padua)
- INRIA Xmas Motion Acquisition Sequences (IXMAS) (INRIA)
- InfAR Dataset –Infrared Action Recognition at Different Times Neurocomputing(Chenqiang Gao, Yinhe Du, Jiang Liu, Jing Lv, Luyu Yang, Deyu Meng, Alexander G. Hauptmann)

- JPL First-Person Interaction dataset - 7 types of human activity videos taken from a first-person viewpoint (Michael S. Ryoo, JPL)
- Jena Action Recognition Dataset - Aibo dog actions (Korner and Denzler)
- K3Da - Kinect 3D Active dataset - K3Da (Kinect 3D active) is a realistic clinically relevant human action dataset containing skeleton, depth data and associated participant information (D. Leightley, M. H. Yap, J. Coulson, Y. Barnouin and J. S. McPhee)
- KIT Robo-Kitchen Activity Data Set - 540 clips of 17 people performing 12 complex kitchen activities.(L. Rybok, S. Friedberger, U. D. Hanebeck, R. Stiefelhagen)
- KTH human action recognition database (KTH CVAP lab)
- Karlsruhe Motion, Intention, and Activity Data set (MINTA) - 7 types of activities of daily living including fully motion primitive segments.(D. Gehrig, P. Krauthausen, L. Rybok, H. Kuehne, U. D. Hanebeck, T. Schultz, R. Stiefelhagen)
- LIRIS Human Activities Dataset - contains (gray/rgb/depth) videos showing people performing various activities (French National Center for Scientific Research)
- LIRIS human activities dataset - 2 cameras, annotated, depth images (Christian Wolf, et al)
- MEXaction2 action detection and localization dataset - to support the development and evaluation of methods for 'spotting' instances of short actions in a relatively large video database: 77 hours, 117 videos (Michel Crucianu and Jenny Benois-Pineau)
- MPII Cooking Activities Dataset (M. Rohrbach)
- MSR-Action3D - benchmark RGB-D action dataset (Microsoft Research Redmond and University of Wollongong)
- MSRActionPair dataset - : Histogram of Oriented 4D Normals for Activity Recognition from Depth Sequences (University of Central Florida and Microsoft)
- MSRC-12 Kinect gesture data set - 594 sequences and 719,359 frames from people performing 12 gestures (Microsoft Research Cambridge)
- MSRC-12 dataset - sequences of human movements, represented as body-part locations, and the associated gesture (Microsoft Research Cambridge and University of Cambridge)
- MSRDailyActivity3D Dataset - There are 16 activities (Microsoft and the Northwestern University)

- ManiAc RGB-D action dataset: different manipulation actions, 15 different versions, 30 different objects manipulated, 20 long and complex chained manipulation sequences (Eren Aksoy)
- Mivia dataset - It consists of 7 high-level actions performed by 14 subjects. (Mivia Lab at the University of Salerno)
- MuHAVi - Multicamera Human Action Video Data (Hossein Ragheb)
- Multi-modal action detection (MAD) Dataset - It contains 35 sequential actions performed by 20 subjects. (CarnegieMellon University)
- Multiview 3D Event dataset - This dataset includes 8 categories of events performed by 8 subjects (University of California at Los Angles)
- NTU RGB+D Action Recognition Dataset - NTU RGB+D is a large scale dataset for human action recognition(Amir Shahroudy)
- Northwestern-UCLA Multiview Action 3D - There are 10 action categories:(Northwestern University and University of California at Los Angles)
- Oxford TV based human interactions (Oxford Visual Geometry Group)
- Parliament - The Parliament dataset is a collection of 228 video sequences, depicting political speeches in the Greek parliament(Michalis Vrigkas, Christophoros Nikou, Ioannins A. kakadiaris)
- RGB-D activity dataset - Each video in the dataset contains 2-7 actions involving interaction with different objects. (Cornell University and Stanford University)
- RGBD-SAR Dataset - RGBD-SAR Dataset (University of Electronic Science and Technology of China and Microsoft)
- Rochester Activities of Daily Living Dataset (Ross Messing)
- SBU Kinect Interaction Dataset - It contains eight types of interactions (Stony Brook University)
- SBU-Kinect-Interaction dataset v2.0 - It comprises of RGB-D video sequences of humans performing interaction activities (Kiwon Yun etc.)
- SDHA Semantic Description of Human Activities 2010 contest - Human Interactions (Michael S. Ryoo, J. K. Aggarwal, Amit K. Roy-Chowdhury)
- SDHA Semantic Description of Human Activities 2010 contest - aerial views (Michael S. Ryoo, J. K. Aggarwal, Amit K. Roy-Chowdhury)
- SFU Volleyball Group Activity Recognition - 2 levels annotations dataset (9 players' actions and 8 scene's activity) for volleyball videos.(M. Ibrahim, S. Muralidharan, Z. Deng, A. Vahdat, and G. Mori / Simon Fraser University)

- [SYSU 3D Human-Object Interaction Dataset](#) - Forty subjects perform 12 distinct activities (Sun Yat-sen University)
- [ShakeFive Dataset](#) - contains only two actions, namely hand shake and high five. (Universiteit Utrecht)
- [ShakeFive2](#) - A dyadic human interaction dataset with limb level annotations on 8 classes in 153 HD videos(Coert van Gemeren, Ronald Poppe, Remco Veltkamp)
- [Sports Videos in the Wild (SVW)](#) - SVW is comprised of 4200 videos captured solely with smartphones by users of Coach Eye smartphone app, a leading app for sports training developed by TechSmith corporation.(Seyed Morteza Safdarnejad, Xiaoming Liu)
- [Stanford Sport Events dataset](#) (Jia Li)
- [THUMOS](#) - Action Recognition in Temporally Untrimmed Videos! - 430 hours of video data and 45 million frames (Gorban, Idrees, Jiang, Zamir, Laptev Shah, Sukthanka)
- [TUM Kitchen Data Set of Everyday Manipulation Activities](#) (Moritz Tenorth, Jan Bandouch)
- [TV Human Interaction Dataset](#) (Alonso Patron-Perez)
- [The Falling Detection dataset](#) - Six subjects in two sceneries performed a series of actions continuously (University of Texas)
- [The TJU dataset](#) - contains 22 actions performed by 20 subjects in two different environments; a total of 1760 sequences. (Tianjin University)
- [The UPCV action dataset](#) - The dataset consists of 10 actions performed by 20 subjects twice. (University of Patras)
- [UC-3D Motion Database](#) - Available data types encompass high resolution Motion Capture, acquired with MVN Suit from Xsens and Microsoft Kinect RGB and depth images.(Institute of Systems and Robotics, Coimbra, Portugal)
- [UCF 101 action dataset](#) 101 action classes, over 13k clips and 27 hours of video data (Univ of Central Florida)
- [UCFKinect](#) - The dataset is composed of 16 actions (University of Central Florida Orlando)
- [UCR Videoweb Multi-camera Wide-Area Activities Dataset](#) (Amit K. Roy-Chowdhury)
- [UTD-MHAD](#) - Eight subjects performed 27 actions four times. (University of Texas at Dallas)

- UTKinect dataset - Ten types of human actions were performed twice by 10 subjects (University of Texas)
- UWA3D Multiview Activity Dataset - Thirty activities were performed by 10 individuals (University of Western Australia)
- Univ of Central Florida - 50 Action Category Recognition in Realistic Videos (3 GB) (Kishore Reddy)
- Univ of Central Florida - ARGAerial camera, Rooftop camera and Ground camera (UCF Computer Vision Lab)
- Univ of Central Florida - Feature Films Action Dataset (Univ of Central Florida)
- Univ of Central Florida - Sports Action Dataset (Univ of Central Florida)
- Univ of Central Florida - YouTube Action Dataset (sports) (Univ of Central Florida)
- Utrecht Multi-Person Motion Benchmark (UMPM). - a collection of video recordings of people together with a ground truth based on motion capture data.(N.P. van der Aa, X. Luo, G.J. Giezeman, R.T. Tan, R.C. Veltkamp.)
- VIRAT Video Dataset - event recognition from two broad categories of activities (single-object and two-objects) which involve both human and vehicles. (Sangmin Oh et al)
- Verona Social interaction dataset (Marco Cristani)
- ViHASi: Virtual Human Action Silhouette Data (userID: VIHASI password: virtual$virtual) (Hossein Ragheb, Kingston University)
- Videoweb (multicamera) Activities Dataset (B. Bhanu, G. Denina, C. Ding, A. Ivers, A. Kamal, C. Ravishankar, A. Roy-Chowdhury, B. Varda)
- WVU Multi-view action recognition dataset (Univ. of West Virginia)
- WorkoutSU-10 Kinect dataset for exercise actions (Ceyhun Akgul)
- WorkoutSU-10 dataset - contains exercise actions selected by professional trainers for therapeutic purposes.(Sabanc University)
- Wrist-mounted camera video dataset - object manipulation (Ohnishi, Kanehira, Kanezaki, Harada)
- YouCook - 88 open-source YouTube cooking videos with annotations (Jason Corso)
- YouTube-8M Dataset -A Large and Diverse Labeled Video Dataset for Video Understanding Research(Google Inc.)

2、Attribute recognition

- [Birds](#)This database contains 600 images (100 samples each) of six different classes of birds.(Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce)
- [Butterflies](#)This database contains 619 images of seven different classes of butterflies. (Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce)
- [CALVIN research group datasets](#) - object detection with eye tracking, imagenet bounding boxes, synchronised activities, stickman and body poses, youtube objects, faces, horses, toys, visual attributes, shape classes (CALVIN ggroup)
- [CelebA](#) - Large-scale CelebFaces Attributes Dataset(Ziwei Liu, Ping Luo, Xiaogang Wang, Xiaoou Tang)
- [HAT](#) database of 27 human attributes (Gaurav Sharma, Frederic Jurie)
- [LFW-10 dataset for learning relative attributes](#) - A dataset of 10,000 pairs of face images with instance-level annotations for 10 attributes.(CVIT, IIIT Hyderabad.)
- [Person Recognition in Personal Photo Collections](#) - we introduced three harder splits for evaluation and long-term attribute annotations and per-photo timestamp metadata.(Oh, Seong Joon and Benenson, Rodrigo and Fritz, Mario and Schiele, Bernt)
- [Visual Attributes Dataset](#) visual attribute annotations for over 500 object classes (animate and inanimate) which are all represented in ImageNet. Each object class is annotated with visual attributes based on a taxonomy of 636 attributes (e.g., has fur, made of metal, is round).
- [WIDER Attribute Dataset](#) - WIDER Attribute is a large-scale human attribute dataset, with 13789 images belonging to 30 scene categories, and 57524 human bounding boxes each annotated with 14 binary attributes.(Li, Yining and Huang, Chen and Loy, Chen Change and Tang, Xiaoou)
- [Animals with Attributes](#) – 30,475 images of 50 animals classes with 6 pre-extracted feature representations for each image.
- [aYahoo and aPascal](#) – Attribute annotations for images collected from Yahoo and Pascal VOC 2008.
- [FaceTracer](#) – 15,000 faces annotated with 10 attributes and fiducial points.
- [PubFig](#) – 58,797 face images of 200 people with 73 attribute classifier outputs.

- LFW – 13,233 face images of 5,749 people with 73 attribute classifier outputs.
- Human Attributes – 8,000 people with annotated attributes. Check also this link for another dataset of human attributes.
- SUN Attribute Database – Large-scale scene attribute database with a taxonomy of 102 attributes.
- ImageNet Attributes – Variety of attribute labels for the ImageNet dataset.
- Relative attributes – Data for OSR and a subset of PubFig datasets. Check also this link for the WhittleSearch data.
- Attribute Discovery Dataset – Images of shopping categories associated with textual descriptions.

3、Autonomous Driving

- AMUSE -The automotive multi-sensor (AMUSE) dataset taken in real traffic scenes during multiple test drives. (Philipp Koschorrek etc.)
- Autonomous Driving - semantic segmentation,pedestrian detection,virtual-world data,far infrared,stereo,driver monitoring(CVC research center and the UAB and UPC universities)
- Joint Attention in Autonomous Driving (JAAD) - The dataset includes instances of pedestrians and cars intended primarily for the purpose of behavioural studies and detection in the context of autonomous driving.(Iuliia Kotseruba, Amir Rasouli and John K. Tsotsos)
- LISA Vehicle Detection Dataset - colour first person driving video under various lighting and traffic conditions (Sivaraman, Trivedi)
- Lost and Found Dataset - The Lost and Found Dataset addresses the problem of detecting unexpected small road hazards (often caused by lost cargo) for autonomous driving applications. (Sebastian Ramos, Peter Pinggera, Stefan Gehrig, Uwe Franke, Rudolf Mester, Carsten Rother)
- SYNTHIA - Large set (~half million) of virtual-world images for training autonomous cars to see. (ADAS Group at Computer Vision Center)
- The SYNTHetic collection of Imagery and Annotations - the purpose of aiding semantic segmentation and related scene understanding problems in the context of driving scenarios (Computer vision center,UAB)

4、Biological/Medical

- [2008 MICCAI MS Lesion Segmentation Challenge](#) (National Institutes of Health Blueprint for Neuroscience Research)
- [ASU DR-AutoCC Data](#) - a Multiple-Instance Learning feature space for a diabetic retinopathy classification dataset (Ragav Venkatesan, Parag Chandakkar, Baoxin Li - Arizona State University)
- [Annotated Spine CT Database](#) for Benchmarking of Vertebrae Localization, 125 patients, 242 scans (Ben Glockern)
- [BRATS](#) - the identification and segmentation of tumor structures in multiparametric magnetic resonance images of the brain (TU Munchen etc.)
- [CRCHistoPhenotypes - Labeled Cell Nuclei Data](#) - colorectal cancer?histology images?consisting of nearly 30,000 dotted nuclei with over 22,000 labeled with the cell type (Rajpoot + Sirinukunwattana)
- [CREMI: MICCAI 2016 Challenge](#) - 6 volumes of electron microscopy of neural tissue,neuron and synapse segmentation, synaptic partner annotation. (Jan Funke, Stephan Saalfeld, Srini Turaga, Davi Bock, Eric Perlman)
- [Cavy Action Dataset](#) - 16 sequences with 640 x 480 resolutions recorded at 7.5 frames per second (fps) with approximately 31621506 frames in total (272 GB) of interacting cavies (guinea pig) (Al-Raziqi and Denzler)
- [Cell Tracking Challenge Datasets](#) - 2D/3D time-lapse video sequences with ground truth(Ma et al., Bioinformatics 30:1609-1617, 2014)
- [Computed Tomography Emphysema Database](#) (Lauge Sorensen)
- [DIADEM: Digital Reconstruction of Axonal and Dendritic Morphology Competition](#) (Allen Institute for Brain Science et al)
- [DIARETDB1 - Standard Diabetic Retinopathy Database](#) (Lappeenranta Univ of Technology)
- [DRIVE: Digital Retinal Images for Vessel Extraction](#) (Univ of Utrecht)
- [DeformIt 2.0](#) - Image Data Augmentation Tool: Simulate novel images with ground truth segmentations from a single image-segmentation pair (Brian Booth and Ghassan Hamarneh)
- [Dermoscopy images](#) (Eric Ehrsam)
- [EPT29.](#)This database contains 4842 images of 1613 specimens of 29 taxa of EPTs:(Tom etc.)

- [FIRE Fundus Image Registration Dataset](#) - 134 retinal image pairs and groud truth for registration.(FORTH‑ICS)
- [KID](#) - A capsule endoscopy database for medical decision support (Anastasios Koulaouzidis and Dimitris Iakovidis)
- [Leaf Segmentation Challenge](#)Tobacco and arabidopsis plant images (Hanno Scharr, Massimo Minervini, Andreas Fischbach, Sotirios A. Tsaftaris)
- [MIT CBCL Automated Mouse Behavior Recognition datasets](#) (Nicholas Edelman)
- [MUCIC: Masaryk University Cell Image Collection](#) - 2D/3D synthetic images of cells/tissues for benchmarking(Masaryk University)
- [MiniMammographic Database](#) (Mammographic Image Analysis Society)
- [Moth fine‑grained recognition](#) - 675 similar classes, 5344 images (Erik Rodner et al)
- [Mouse Embryo Tracking Database](#) - cell division event detection (Marcelo Cicconet, Kris Gunsalus)
- [OASIS](#) - Open Access Series of Imaging Studies - 500+ MRI data sets of the brain (Washington University, Harvard University, Biomedical Informatics Research Network)
- [Plant Phenotyping Datasets](#) - plant data suitable for plant and leaf detection, segmentation, tracking, and species recognition (M. Minervini, A. Fischbach, H. Scharr, S. A. Tsaftaris)
- [Retinal fundus images - Ground truth of vascular bifurcations and crossovers](#) (Univ of Groningen)
- [Spine and Cardiac data](#) (Digital Imaging Group of London Ontario, Shuo Li)
- [Stonefly9](#)This database contains 3826 images of 773 specimens of 9 taxa of Stoneflies (Tom etc.)
- [Univ of Central Florida - DDSM: Digital Database for Screening Mammography](#) (Univ of Central Florida)
- [VascuSynth](#) - 120 3D vascular tree like structures with ground truth (Mengliu Zhao, Ghassan Hamarneh)
- [VascuSynth](#) - Vascular Synthesizer generates vascular trees in 3D volumes. (Ghassan Hamarneh, Preet Jassi, Mengliu Zhao)
- [York Cardiac MRI dataset](#) (Alexander Andreopoulos)
- [Caltech-UCSD Birds Dataset](#) – Hundreds of bird categories with annotated parts and attributes.

- Stanford Dogs Dataset – 20,000 images of 120 breeds of dogs from around the world.
- Oxford-IIIT Pet Dataset – 37 category pet dataset with roughly 200 images for each class. Pixel level trimap segmentation is included.
- Leeds Butterfly Dataset – 832 images of 10 species of butterflies.
- Oxford Flower Dataset – Hundreds of flower categories.

5、Camera calibration

- Catadioptric camera calibration images (Yalin Bastanlar)
- GoPro-Gyro Dataset - This dataset consists of a number of wide-angle rolling shutter video sequences with corresponding gyroscope measurements (Hannes etc.)
- LO-RANSAC - LO-RANSAC library for estimation of homography and epipolar geometry(K. Lebeda, J. Matas and O. Chum)

6、Face and Eye/Iris Databases

- 300 Videos in the Wild (300-VW) - 68 Facial Landmark Tracking (Chrysos, Antonakos, Zafeiriou, Snape, Shen, Kossaifi, Tzimiropoulos, Pantic)
- 3D Mask Attack Database (3DMAD) - 76500 frames of 17 persons using Kinect RGBD with eye positions (Sebastien Marcel)
- 3D facial expression - Binghamton University 3D Static and Dynamic Facial Expression Databases (Lijun Yin, Jeff Cohn, and teammates)
- Audio-visual database for face and speaker recognition (Mobile Biometry MOBIO http://www.mobioproject.org/)
- BANCA face and voice database (Univ of Surrey)
- Binghampton Univ 3D static and dynamic facial expression database (Lijun Yin, Peter Gerhardstein and teammates)
- BioID face database (BioID group)
- BioVid Heat Pain Database - This video (and biomedical signal) dataset contains facial and physiopsychological reactions of 87 study participants who were subjected to experimentally induced heat pain.(University of Magdeburg (Neuro-Information Technology group) and University of Ulm (Emotion Lab))

- [Biwi 3D Audiovisual Corpus of Affective Communication](#) - 1000 high quality, dynamic 3D scans of faces, recorded while pronouncing a set of English sentences.
- [Bosphorus 3D/2D Database of FACS annotated facial expressions, of head poses and of face occlusions](#) (Bogazici University)
- [CASIA-IrisV3](#) (Chinese Academy of Sciences, T. N. Tan, Z. Sun)
- [CASIR Gaze Estimation Database](#) - RGB and depth images (from Kinect V1.0) and ground truth values of facial features corresponding to experiments for gaze estimation benchmarking: (Filipe Ferreira etc.)
- [CMU Facial Expression Database](#) (CMU/MIT)
- [CMU Pose, Illumination, and Expression (PIE) Database](#) (Simon Baker)
- [CMU/MIT Frontal Faces](#) (CMU/MIT)
- [CMU/MIT Frontal Faces](#) (CMU/MIT)
- [CSSE Frontal intensity and range images of faces](#) (Ajmal Mian)
- [CelebA](#) - Large-scale CelebFaces Attributes Dataset(Ziwei Liu, Ping Luo, Xiaogang Wang, Xiaoou Tang)
- [Cohn-Kanade AU-Coded Expression Database](#) - 500+ expression sequences of 100+ subjects, coded by activated Action Units (Affect Analysis Group, Univ. of Pittsburgh)
- [Columbia Gaze Data Set](#) - 5,880 images of 56 people over 5 head poses and 21 gaze directions (Brian A. Smith, Qi Yin, Steven K. Feiner, Shree K. Nayar)
- [Computer Vision Laboratory Face Database (CVL Face Database)](#) - Database contains 798 images of 114 persons, with 7 images per person and is freely available for research purposes.(Peter Peer etc.)
- **Deng Cai's face dataset in Matlab Format** – Easy to use if you want play with simple face datasets including Yale, ORL, PIE, and Extended Yale B.
- [EURECOM Facial Cosmetics Database](#) - 389 images, 50 persons with/without make-up, annotations about the amount and location of applied makeup.(Jean-Luc DUGELAY et al)
- [EURECOM Kinect Face Database](#) - 52 people, 2 sessions, 9 variations, 6 facial landmarks.(Jean-Luc DUGELAY et al)
- [EYEDIAP dataset](#) - The EYEDIAP dataset was designed to train and evaluate gaze estimation algorithms from RGB and RGB-D data.It contains a

diversity of participants, head poses, gaze targets and sensing conditions.(Kenneth Funes and Jean-Marc Odobez)

- FDDB: Face Detection Data set and Benchmark - studying unconstrained face detection (University of Massachusetts Computer Vision Laboratory)
- FG-Net Aging Database of faces at different ages (Face and Gesture Recognition Research Network)
- Face Recognition Grand Challenge datasets (FRVT - Face Recognition Vendor Test)
- Face Super-Resolution Dataset - Ground truth HR-LR face images captured with a dual-camera setup(Chengchao Qu etc.)
- FaceScrub - A Dataset With Over 100,000 Face Images of 530 People (50:50 male and female) (H.-W. Ng, S. Winkler)
- FaceTracer Database - 15,000 faces (Neeraj Kumar, P. N. Belhumeur, and S. K. Nayar)
- Facial Recognition Technology (FERET) Database (USA National Institute of Standards and Technology)
- Hannah and her sisters database - a dense audio-visual person-oriented ground-truth annotation of faces, speech segments, shot boundaries (Patrick Perez, Technicolor)
- Hong Kong Face Sketch Database
- IDIAP Head Pose Database (IHPD) - The dataset contains a set of meeting videos along with the head groundtruth of individual participants (around 128min)(Sileye Ba and Jean-Marc Odobez)
- IMDB-WIKI - 500k+ face images with age and gender labels (Rasmus Rothe, Radu Timofte, Luc Van Gool)
- Iranian Face Database - IFDB is the first image database in middle-east, contains color facial images with age, pose, and expression whose subjects are in the range of 2-85. (Mohammad Mahdi Dehshibi)
- Japanese Female Facial Expression (JAFFE) Database (Michael J. Lyons)
- LFW: Labeled Faces in the Wild - unconstrained face recognition
- MIT CBCL Face Recognition Database (Center for Biological and Computational Learning)
- MIT Collation of Face Databases (Ethan Meyers)
- MIT eye tracking database (1003 images) (Judd et al)

- [MMI Facial Expression Database](#) - 2900 videos and high-resolution still images of 75 subjects, annotated for FACS AUs.
- [MORPH (Craniofacial Longitudinal Morphological Face Database)](#) (University of North Carolina Wilmington)
- [Manchester Annotated Talking Face Video Dataset](#) (Timothy Cootes)
- [MegaFace](#) - 1 million faces in bounding boxes (Kemelmacher-Shlizerman, Seitz, Nech, Miller, Brossard)
- [NIST Face Recognition Grand Challenge (FRGC)](#) (NIST)
- [NIST mugshot identification database](#) (USA National Institute of Standards and Technology)
- [Notre Dame Iris Image Dataset](#) (Patrick J. Flynn)
- [Notre Dame face, IR face, 3D face, expression, crowd, and eye biometric datasets](#) (Notre Dame)
- [ORL face database: 40 people with 10 views](#) (ATT Cambridge Labs)
- [OUI-Adience Faces](#) - unfiltered faces for gender and age classification plus 3D faces (OUI)
- [Oxford: faces, flowers, multi-view, buildings, object categories, motion segmentation, affine covariant regions, misc](#) (Oxford Visual Geometry Group)
- [PubFig: Public Figures Face Database](#) (Neeraj Kumar, Alexander C. Berg, Peter N. Belhumeur, and Shree K. Nayar)
- [Re-labeled Faces in the Wild](#) - original images, but aligned using "deep funneling" method. (University of Massachusetts, Amherst)
- [SCface](#) - Surveillance Cameras Face Database (Mislav Grgic, Kresimir Delac, Sonja Grgic, Bozidar Klimpak))
- [Salient features in gaze-aligned recordings of human visual input](#) - TB of human gaze-contingent data "in the wild" (Frank Schumann etc.)
- [SiblingsDB](#) - The SiblingsDB contains two datasets depicting images of individuals related by sibling relationships. (Politecnico di Torino/Computer Graphics & Vision Group)
- [Trondheim Kinect RGB-D Person Re-identification Dataset](#) (Igor Barros Barbosa)
- [UB KinFace Database](#) - University of Buffalo kinship verification and recognition database
- [UBIRIS: Noisy Visible Wavelength Iris Image Databases](#) (University of Beira)

- UTIRIS cross-spectral iris image databank (Mahdi Hosseini)
- VIPSL Database - VIPSL Database is for research on face sketch-photo synthesis and recognition, including 200 subjects (1 photo and 5 sketches per subject).(Nannan Wang)
- WIDER FACE: A Face Detection Benchmark - 32,203 images with 393,703 labeled faces, 61 event classes (Shuo Yang, Ping Luo, Chen Change Loy, Xiaoou Tang)
- XM2VTS Face video sequences (295): The extended M2VTS Database (XM2VTS) - (Surrey University)
- Yale Face Database - 11 expressions of 10 people (A. Georghaides)
- Yale Face Database B - 576 viewing conditions of 10 people (A. Georghaides)
- York Univ Eye Tracking Dataset (120 images) (Neil Bruce)
- YouTube Faces DB - 3,425 videos of 1,595 different people. (Wolf, Hassner, Maoz)
- Zurich Natural Image - the image material used for creating natural stimuli in a series of eye-tracking studies (Frey et al.)

7、Fingerprints

- FVC fingerpring verification competition 2002 dataset (University of Bologna)
- FVC fingerpring verification competition 2004 dataset (University of Bologna)
- NIST fingerprint databases (USA National Institute of Standards and Technology)
- SPD2010 Fingerprint Singular Points Detection Competition (SPD 2010 committee)

8、General Images

- A Dataset for Real Low-Light Image Noise Reduction - It contains pixel and intensity aligned pairs of images corrupted by low-light camera noise and their low-noise counterparts. (J. Anaya, A. Barbu)
- AMOS: Archive of Many Outdoor Scenes (20+m) (Nathan Jacobs)

- [Aerial images](#) Building detection from aerial images using invariant color features and shadow information (Beril Sirmacek)
- [BGU Hyperspectral Image Database of Natural Scenes](#) (Ohad Ben-Shahar and Boaz Arad)
- [Brown Univ Large Binary Image Database](#) (Ben Kimia)
- [CMP Facade Database](#) - Includes 606 rectified images of facades from various places with 12 architectural classes annotated.(Radim Tylecek)
- [Caltech-UCSD Birds-200-2011](#) (Catherine Wah)
- [Columbia Multispectral Image Database](#) (F. Yasuma, T. Mitsunaga, D. Iso, and S.K. Nayar)
- [DAQUAR (Visual Turing Challenge)](#) - A dataset containing questions and answers about real-world indoor scenes.(Mateusz Malinowski, Mario Fritz)
- [Dataset of American Movie Trailers 2010-2014](#) - Contains links to 474 hollywood movie trailers along with associated metadata (genre, budget, runtime, release, MPAA rating, screens released, sequel indicator) (USC Signal Analysis and Interpretation Lab)
- [General 100 Dataset](#) - General-100 dataset contains 100 bmp-format images (with no compression), which are well-suited for super-resolution training(Dong, Chao and Loy, Chen Change and Tang, Xiaoou)
- [HIPR2 Image Catalogue of different types of images](#) (Bob Fisher et al)
- [Hyperspectral images for spatial distributions of local illumination in natural scenes](#) - Thirty calibrated hyperspectral radiance images of natural scenes with probe spheres embedded for local illumination estimation. (Nascimento, Amano & Foster)
- [Hyperspectral images of natural scenes - 2002](#) (David H. Foster)
- [Hyperspectral images of natural scenes - 2004](#) (David H. Foster)
- [ISPRS multi-platform photogrammetry dataset](#) - 1) Nadir and oblique aerial images plus 2) Combined UAV and terrestrial images (Francesco Nex and Markus Gerke)
- [ImageNet Large Scale Visual Recognition Challenges](#) - currently 200 object classes and 500+K images (Alex Berg, Jia Deng, Fei-Fei Li and others)
- [ImageNet Linguistically organised (WordNet) Hierarchical Image Database - 10E7 images, 15K categories](#) (Li Fei-Fei, Jia Deng, Hao Su, Kai Li)
- [Improved 3D Sparse Maps for High-performance Structure from Motion with Low-cost Omnidirectional Robots - Evaluation Dataset](#) - Data set used

in research paper doi:10.1109/ICIP.2015.7351744 (Breckon, Toby P., Cavestany, Pedro)

- LabelMeFacade Database - 945 labeled building images (Erik Rodner et al)
- Local illumination hyperspectral radiance images - Thirty hyperspectral radiance images of natural scenes with embedded probe spheres for local illumination estimates(Sgio M. C. Nascimento, Kinjiro Amano, David H. Foster)
- McGill Calibrated Colour Image Database (Adriana Olmos and Fred Kingdom)
- MIT LabelMe – Online annotation tool for building computer vision databases.
- Multiply Distorted Image Database -a database for evaluating the results of image quality assessment metrics on multiply distorted images.(Fei Zhou)
- NPRgeneral - A standardized collection of images for evaluating image stylization algorithms. (David Mould, Paul Rosin)
- NYU Symmetry Database - 176 single-symmetry and 63 multyple-symmetry images (Marcelo Cicconet and Davi Geiger)
- OTCBVS Thermal Imagery Benchmark Dataset Collection (Ohio State Team)
- PAnorama Sparsely STructured Areas Datasets - the PASSTA datasets used for evaluation of the image alignment (Andreas Robinson)
- Time-Lapse Hyperspectral Radiance Images of Natural Scenes - Four time-lapse sequences of 7-9 calibrated hyperspectral radiance images of natural scenes taken over the day. (Foster, D.H., Amano, K., & Nascimento, S.M.C.)
- Time-lapse hyperspectral radiance images - Four time-lapse sequences of 7-9 calibrated hyperspectral images of natural scenes, spectra at 10-nm intervals(David H. Foster, Kinjiro Amano, Sgio M. C. Nascimento)
- Tiny Images Dataset 79 million 32x32 color images (Fergus, Torralba, Freeman)
- Visual Question Answering - 254K imags, 764K questions, ground truth (Agrawal, Lu, Antol, Mitchell, Zitnick, Batra, Parikh)
- Visual Question Generation - 15k images (including both object-centric and event-centric images), 75k natural questions asked about the images which can evoke further conversation(Nasrin Mostafazadeh , Ishan Misra , Jacob Devlin , Margaret Mitchell , Xiao dong He , Lucy Vanderwende)

- YFCC100M: The New Data in Multimedia Research - This publicly available curated dataset of 100 million photos and videos is free and legal for all.(Bart ThomeeYahoo Labs and Flickr in San Francisco,etc.)

9、General RGBD and Depth Datasets

Note: there are 3D datasets elsewhere as well, *e.g.* in Objects, Scenes, and Actions.

- 3D-Printed RGB-D Object Dataset - 5 objects with groundtruth CAD models and camera trajectories, recorded with various quality RGB-D sensors(Siemens & TUM)
- 3DCOMET - 3DCOMET is a dataset for testing 3D data compression methods.(Miguel Cazorla , Javier Navarrete,Vicente Morell, Miguel Cazorla, Diego Viejo, Jose Garcia-Rodriguez, Sergio Orts.)
- A Dataset for Non-Rigid Reconstruction from RGB-D Data - Eight scenes for reconstructing non-rigid geometry from RGB-D data, each containing several hundred frames along with our results. (Matthias Innmann, Michael Zollhoefer, Matthias Niessner, Christian Theobalt, Marc Stamminger)
- A Large Dataset of Object Scans - 392 objects in 9 casses, hundreds of frames each (Choi, Zhou, Miller, Koltun)
- Articulated Object Challenge - 4 articulated objects consisting of rigids parts connected by 1D revolute and prismatic joints, 7000+ RGBD images with annotations for 6D pose estimation(Frank Michel, Alexander Krull, Eric Brachmann, Michael. Y. Yang,Stefan Gumhold, Carsten Rother)
- BigBIRD - 100 objects with for each object, 600 3D point clouds and 600 high-resolution color images spanning all views (Singh, Sha, Narayan, Achim, Abbeel)
- CAESAR Civilian American and European Surface Anthropometry Resource Project - 4000 3D human body scans (SAE International)
- CIN 2D+3D object classification dataset - segmented color and depth images of objects from 18 categories of common household and office objects (Bj?rn Browatzki et al)
- CTU Garment Folding Photo Dataset - Color and depth images from various stages of garment folding.(Sushkov R., Melkumov I., Smutn y V. (Czech Technical University in Prague))

- CTU Garment Sorting Dataset - Dataset of garment images, detailed stereo images, depth images and weights.(Petrik V., Wagner L. (Czech Technical University in Prague))
- Clothing part dataset - The clothing part dataset consists of image and depth scans, acquired with a Kinect, of garments laying on a table, with over a thousand part annotations (collar, cuffs, hood, etc) using polygonal masks.(Arnau Ramisa, Guillem Aleny,Francesc Moreno-Noguer and Carme Torras)
- Cornell-RGBD-Dataset - Office Scenes (Hema Koppula)
- Delft Windmill Interior and Exterior Laser Scanning Point Clouds -(Beril Sirmacek)
- EURECOM Kinect Face Database - 52 people, 2 sessions, 9 variations, 6 facial landmarks.(Jean-Luc DUGELAY et al)
- EURECOM Kinect Face Database - 52 people, 2 sessions, 9 variations, 6 facial landmarks.(Jean-Luc DUGELAY et al)
- Goldfinch: GOogLe image-search Dataset for FINe grained CHallenges - a largescale dataset for finegrained bird (11K species),butterfly (14K species), aircraft (409 types), and dog (515 breeds) recognition.(Jonathan Krause, Benjamin Sapp,Andrew Howard, Howard Zhou, Alexander Toshev, Tom Duerig, James Philbin, Li Fei-Fei)
- IMPART multi-view/multi-modal 2D+3D film production dataset - LIDAR, video, 3D models, spherical camera, RGBD, stereo, action, facial expressions, etc. (Univ. of Surrey)
- Kinect v2 Dataset - Efficient Multi-Frequency Phase Unwrapping using Kernel Density Estimation (Felix etc.)
- Multi-sensor 3D Object Dataset for Object Recognition with Full Pose Estimation - Multi-sensor 3D Object Dataset for Object Recognition and Pose Estimation(Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea,etc.)
- NTU RGB+D Action Recognition Dataset - NTU RGB+D is a large scale dataset for human action recognition(Amir Shahroudy)
- NYU Depth Dataset V2 - Indoor Segmentation and Support Inference from RGBD Images
- Oakland 3-D Point Cloud Dataset (Nicolas Vandapel)

- [SYNTHIA](#) - Large set (~half million) of virtual-world images for training autonomous cars to see. (ADAS Group at Computer Vision Center)
- [Semantic-8](#): 3D point cloud classification with 8 classes (ETH Zurich)
- [Stereo and ToF dataset with ground truth](#) - The dataset contains 5 different scenes acquired with a Time-of-flight sensor and a stereo setup. Ground truth information is also provided.(Carlo Dal Mutto, Pietro Zanuttigh, Guido M. Cortelazzo)
- [TUM RGB-D Benchmark](#) - Dataset and benchmark for the evaluation of RGB-D visual odometry and SLAM algorithms (BCrgen Sturm, Nikolas Engelhard, Felix Endres, Wolfram Burgard and Daniel Cremers)
- [UC-3D Motion Database](#) - Available data types encompass high resolution Motion Capture, acquired with MVN Suit from Xsens and Microsoft Kinect RGB and depth images.(Institute of Systems and Robotics, Coimbra, Portugal)
- [Washington RGB-D Object Dataset](#) - 300 common household objects and 14 scenes. (University of Washington and Intel Labs Seattle)

10、General Videos

- [AlignMNIST](#) - An artificially extended version of the MNIST handwritten dataset. (en Hauberg)
- [GoPro-Gyro Dataset](#) - ego centric videos (Linkoping Computer Vision Laboratory)
- [Large scale YouTube video dataset](#) - 156,823 videos (2,907,447 keyframes) crawled from YouTube videos (Yi Yang)
- [MovieQA](#) - each machines to understand stories by answering questions about them. 15000 multiple choice QAs, 400+ movies.(M. Tapaswi, Y. Zhu, R. Stiefelhagen, A. Torralba, R. Urtasun, and S. Fidler)
- [Multispectral visible-NIR video sequences](#) - Annotated multispectral video, visible + NIR (LE2I, Universit de Bourgogne)
- [Sports-1M](#) - Dataset for sports video classification containing 487 classes and 1.2M videos.(Andrej Karpathy and George Toderici and Sanketh Shetty and Thomas Leung and Rahul Sukthankar and Li Fei-Fei.)

- [Video Object Segmentation dataset DAVIS](#) - Densely Annotated VIdeo Segmentation (F. Perazzi, J. Pont-Tuset, B. McWilliams, L. Van Gool, M. Gross, and A. Sorkine-Hornung)
- [Video Sequences](#)used for research on Euclidean upgrades based on minimal assumptions about the camera(Kenton McHenry)
- [Video Stacking Dataset](#) - A Virtual Tripod for Hand-held Video Stacking on Smartphones (Erik Ringaby etc.)
- [YFCC100M videos](#) - A benchmark on the video subset of YFCC100M which includes the videos, he video content features and the API to a sate-of-the-art video content engine.(Lu Jiang)
- [YFCC100M: The New Data in Multimedia Research](#) - This publicly available curated dataset of 100 million photos and videos is free and legal for all.(Bart ThomeeYahoo Labs and Flickr in San Francisco,etc.)
- [YouTube-8M](#) - Dataset for video classification in the wild, containing pre-extracted frame level features from 8M videos, and 4800 classes.(Sami Abu-El-Haija, Nisarg Kothari, Joonseok Lee, Paul Natsev,George Toderici, Balakrishnan Varadarajan, Sudheendra Vijayanarasimhan)
- [YouTube-8M Dataset](#) -A Large and Diverse Labeled Video Dataset for Video Understanding Research(Google Inc.)

11、Hand, Hand Grasp, Hand Action and Gesture Databases

- [3D Articulated Hand Pose Estimation with Single Depth Images](#) (Tang, Chang, Tejani, Kim, Yu)
- [A Dataset of Human Manipulation Actions](#) - RGB-D of 25 objects and 6 actions (Alessandro Pieropan)
- [A Hand Gesture Detection Dataset](#) (Javier Molina et al)
- [A-STAR Annotated Hand-Depth Image Dataset and its Performance Evaluation](#) - depth data and data glove data, 29 images of 30 volunteers, Chinese number counting and American Sign Language (Xu and Cheng)
- [Bosphorus Hand Geometry Database and Hand-Vein Database](#) (Bogazici University)
- [EgoHands](#) - A large dataset with over 15,000 pixel-level-segmented hands recorded from egocentric cameras of people interacting with each other. (Sven Bambach)

- [FORTH Hand tracking library](#) (FORTH)
- [General HANDS: general hand detection and pose challenge](#) - 22 sequences with different gestures, activities and viewpoints (UC Irvine)
- [Grasp UNderstanding (GUN-71) dataset](#) - 12,000 first-person RGB-D images of object manipulation scenes annotated using a taxonomy of 71 fine-grained grasps.(Rogez, Supancic and Ramanan)
- [Hand gesture and marine silhouettes](#) (Euripides G.M. Petrakis)
- [HandNet: annotated depth images of articulated hands](#) 214971 annotated depth images of hands captured by a RealSense RGBD sensor of hand poses. Annotations: per pixel classes, 6D fingertip pose, heatmap. Images -> Train: 202198, Test: 10000, Validation: 2773. Recorded at GIP Lab, Technion.
- [IDIAP Hand pose/gesture datasets](#) (Sebastien Marcel)
- [Kinect and Leap motion gesture recognition dataset](#) - The dataset contains 1400 different gestures acquired with both the Leap Motion and the Kinect devices(Giulio Marin, Fabio Dominio, Pietro Zanuttigh)
- [Kinect and Leap motion gesture recognition dataset](#) - The dataset contains several different static gestures acquired with the Creative Senz3D camera.(A. Memo, L. Minto, P. Zanuttigh)
- [LISA CVRR-HANDS 3D](#) - 19 gestures performed by 8 subjects as car driver and passengers (Ohn-Bar and Trivedi)
- [LISA Vehicle Detection Dataset](#) - colour first person driving video under various lighting and traffic conditions (Sivaraman, Trivedi)
- [MPI Dexter 1 Dataset for Evaluation of 3D Articulated Hand Motion Tracking](#) - Dexter 1: 7 sequences of challenging, slow and fast hand motions, RGB + depth (Sridhar, Oulasvirta, Theobalt)
- [MSR Realtime and Robust Hand Tracking from Depth](#) - (Qian, Sun, Wei, Tang, Sun)
- [Mobile and Webcam Hand images database](#) - MOHI and WEHI - 200 people, 30 images each (Ahmad Hassanat)
- [NYU Hand Pose Dataset](#) - 8252 test-set and 72757 training-set frames of captured RGBD data with ground-truth hand-pose, 3 views (Tompson, Stein, Lecun, Perlin}

- Sahand Dynamic Hand Gesture Database - This database contains 11 Dynamic gestures designed to convey the functions of mouse and touch screens to computers.(Behnam Maleki, Hossein Ebrahimnezhad)
- Sheffield gesture database - 2160 RGBD hand gesture sequences, 6 subjects, 10 gestures, 3 postures, 3 backgrounds, 2 illuminations (Ling Shao)
- UT Grasp Data Set - 4 subjects grasping a variety of objectss with a variety of grasps (Cai, Kitani, Sato)
- Yale human grasping data set - 27 hours of video with tagged grasp, object, and task data from two housekeepers and two machinists (Bullock, Feix, Dollar)

12、Image, Video and Shape Database Retrieval

- ANN_SIFT1M - 1M Flickr images encoded by 128D SIFT descriptors (Jegou et al)
- Brown Univ 25/99/216 Shape Databases (Ben Kimia)
- CIFAR-10 - 60K 32x32 images from 10 classes, with a 512D GIST descriptor (Alex Krizhevsky)
- CLEF-IP 2011 evaluation on patent images
- Dataset of Structured Queries and Spatial Relations - Dataset of structured queries about images with the emphasise on spatial relations.(Mateusz Malinowski, Mario Fritz)
- DeepFashion - Large-scale Fashion Database(Ziwei Liu, Ping Luo, Shi Qiu, Xiaogang Wang, Xiaoou Tang)
- EMODB - Thumbnails of images in the picsearch image search engine together with the picsearch emotion keywords (Reiner Lenz etc.)
- ETU10 Silhouette Dataset - The dataset consists of 720 silhouettes of 10 objects, with 72 views per object.(M. Akimaliev and M.F. Demirci)
- Flickr 30K - images, actions and captions (Peter Young et al)
- Flickr15k - Sketch based Image Retrieval (SBIR) Benchmark - Dataset of 330 sketches and 15,024 photos comprising 33 object categories,benchmark dataset commonly used to evaluate Sketch based Image Retrieval (SBIR) algorithms.(Hu and Collomosse, CVIU 2013)
- IAPR TC-12 Image Benchmark (Michael Grubinger)

- IAPR-TC12 Segmented and annotated image benchmark (SAIAPR TC-12): (Hugo Jair Escalante)
- ImageCLEF 2010 Concept Detection and Annotation Task (Stefanie Nowak)
- ImageCLEF 2011 Concept Detection and Annotation Task - multi-label classification challenge in Flickr photos
- METU Trademark datasetThe METU Dataset is composed of more than 900K real logos belonging to companies worldwide. (Usta Bilgi Sistemleri A.S. and Grup Ofis Marka Patent A.S)
- MPI Movie Description dataset - text and video (A. Rohrbach)
- McGill 3D Shape Benchmark (Siddiqi, Zhang, Macrini, Shokoufandeh, Bouix, Dickinson)
- Multiview Stereo Evaluation - Each dataset is registered with a "ground-truth" 3D model acquired via a laser scanning process(Steve Seitz et al)
- NIST SHREC - other NIST retrieval contest databases and links (USA National Institute of Standards and Technology)
- NIST SHREC 2010 - Shape Retrieval Contest of Non-rigid 3D Models (USA National Institute of Standards and Technology)
- NIST TREC Video Retrieval Evaluation Database (USA National Institute of Standards and Technology)
- NUS-WIDE - 269K Flickr images annotated with 81 concept tags, enclded as a 500D BoVW descriptorChau et al)
- Princeton Shape Benchmark (Princeton Shape Retrieval and Analysis Group)
- Queensland cross media dataset - millions of images and text documents for "cross-media" retrieval (Yi Yang)
- TOSCA 3D shape database (Bronstein, Bronstein, Kimmel)
- YouTube-8M Dataset -A Large and Diverse Labeled Video Dataset for Video Understanding Research(Google Inc.)

13、 Object Databases

- 2.5D/3D Datasets of various objects and scenes (Ajmal Mian)
- 3D Object Recognition Stereo DatasetThis dataset consists of 9 objects and 80 test images. (Akash Kushal and Jean Ponce)
- 3D Photography Dataseta collection of ten multiview data sets captured in our lab(Yasutaka Furukawa and Jean Ponce)

- [3D-Printed RGB-D Object Dataset](#) - 5 objects with groundtruth CAD models and camera trajectories, recorded with various quality RGB-D sensors(Siemens & TUM)
- [Amsterdam Library of Object Images (ALOI): 100K views of 1K objects](#) (University of Amsterdam/Intelligent Sensory Information Systems)
- [B3DO: Berkeley 3-D Object Dataset](#) - household object detection (Janoch et al)
- [Beyond PASCAL: A Benchmark for 3D Object Detection in the Wild](#) - 12 class, 3000+ images each with 3D annotations (Yu Xiang, Roozbeh Mottaghi, Silvio Savarese)
- [Bristol Egocentric Object Interactions Dataset](#) - egocentric object interactions with synchronised gaze (Dima Damen)
- [CORE image dataset](#) - to help learn more detailed models and for exploring cross-category generalization in object recognition. (Ali Farhadi, Ian Endres, Derek Hoiem, and David A. Forsyth)
- [CTU Color and Depth Image Dataset of Spread Garments](#) - Images of spread garments with annotated corners.(Wagner, L., Krejov D., and Smutn V. (Czech Technical University in Prague))
- [Caltech 101 (now 256) category object recognition database](#) (Li Fei-Fei, Marco Andreeto, Marc'Aurelio Ranzato)
- [Catania Fish Species Recognition](#) - 15 fish species, with about 20,000 sample training images and additional test images (Concetto Spampinato))
- [Columbia COIL-100 3D object multiple views](#) (Columbia University)
- [Densely sampled object views: 2500 views of 2 objects, eg for view-based recognition and modeling](#) (Gabriele Peters, Universiteit Dortmund)
- [EDUB-Obj](#) - Egocentric dataset for object localization and segmentation.(Marc Bolaños and Petia Radeva.)
- [Ellipse finding dataset](#) (Dilip K. Prasad et al)
- [GDXray:X-ray images for X-ray testing and Computer Vision](#) - GDXray includes five groups of images: Castings, Welds*,Baggages, Nature and Settings. (Domingo Mery, Catholic University of Chile)
- [GRAZ-02 Database (Bikes, cars, people)](#) (A. Pinz)
- [GTSDB: German Traffic Sign Detection Benchmark](#) (Ruhr-Universitat Bochum)

- [ICubWorld](#) – iCubWorld datasets are collections of images acquired by recording from the cameras of the iCub humanoid robot while it observes daily objects.(Giulia Pasquale, Carlo Ciliberto, Giorgio Metta, Lorenzo Natale, Francesca Odone and Lorenzo Rosasco.)
- [LISA Traffic Light Dataset](#) – 6 light classes in various lighting conditions (Jensen, Philipsen, Mogelmose, Moeslund, and Trivedi)
- [LISA Traffic Sign Dataset](#) – video of 47 US sign types with 7855 annotations on 6610 frames (Mogelmose, Trivedi, and Moeslund)
- [Linkoping 3D Object Pose Estimation Database](#) (Fredrik Viksten and Per-Erik Forssen)
- [Linkoping Traffic Signs Dataset](#) – 3488 traffic signs in 20K images (Larsson and Felsberg)
- [MIT CBCL Car Data](#) (Center for Biological and Computational Learning)
- [MIT CBCL StreetScenes Challenge Framework:](#) (Stan Bileschi)
- [Microsoft COCO](#) – Common Objects in Context (Tsung-Yi Lin et al)
- [Microsoft Object Class Recognition image databases](#) (Antonio Criminisi, Pushmeet Kohli, Tom Minka, Carsten Rother, Toby Sharp, Jamie Shotton, John Winn)
- [Microsoft salient object databases (labeled by bounding boxes)](#) (Liu, Sun Zheng, Tang, Shum)
- [ModelNet](#) – 127,915 CAD Models, 662 Object Categories, 10 Categories with Annotated Orientation (Wu, Song, Khosla, Yu, Zhang, Tang, Xiao)
- [NABirds Dataset](#) – 70,000 annotated photographs of the 400 species of birds commonly observed in North America (Grant Van Horn)
- [NEC Toy animal object recognition or categorization database](#) (Hossein Mobahi)
- [NORB 50 toy image database](#) (NYU)
- [NTU-VOI: NTU Video Object Instance Dataset](#) – video clips with frame-level bounding box annotations of object instances for evaluating object instance search and localization in large scale videos.(Jingjing Meng, et. al.)
- [Object Pose Estimation Database](#) – This database contains 16 objects, each sampled at 5 degrees angle increments along two rotational axes (F. Viksten etc.)

- Object Recognition DatabaseThis database features modeling shots of eight objects and 51 cluttered test shots containing multiple objects.(Fred Rothganger, Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce.)
- PASCAL 2007 Challange Image Database (motorbikes, cars, cows) (PASCAL Consortium)
- PASCAL 2008 Challange Image Database (PASCAL Consortium)
- PASCAL 2009 Challange Image Database (PASCAL Consortium)
- PASCAL 2010 Challange Image Database (PASCAL Consortium)
- PASCAL 2011 Challange Image Database (PASCAL Consortium)
- PASCAL 2012 Challange Image Database Category classification, detection, and segmentation, and still-image action classification (PASCAL Consortium)
- PASCAL Image Database (motorbikes, cars, cows) (PASCAL Consortium)
- PASCAL Parts dataset - PASCAL VOC with segmentation annotation for semantic parts of objects (Alan Yuille)
- PASCAL-Context dataset - annotations for 400+ additional categories (Alan Yuille)
- Raindrop Detection - Improved Raindrop Detection using Combined Shape and Saliency Descriptors with Scene Context Isolation - Evaluation Dataset (Breckon, Toby P., Webster, Dereck D.)
- Swedish Leaf Dataset - These images contains leaves from 15 treeclasses (Oskar J. O. S?derkvist)
- UAH Traffic Signs Dataset (Arroyo etc.)
- UIUC Car Image Database (UIUC)
- UIUC Dataset of 3D object categories (S. Savarese and L. Fei-Fei)
- Venezia 3D object-in-clutter recognition and segmentation (Emanuele Rodola)
- Visual Attributes Dataset visual attribute annotations for over 500 object classes (animate and inanimate) which are all represented in ImageNet. Each object class is annotated with visual attributes based on a taxonomy of 636 attributes (e.g., has fur, made of metal, is round).
- Visual Genome: Connecting Language and Vision Using Crowdsourced Dense Image Annotations - Visual Genome is a dataset, a knowledge base, an ongoing effort to connect structured image concepts to language.(Ranjay Krishna, Yuke Zhu, Oliver Groth,Justin Johnson, Kenji

Hata, Joshua Kravitz, Stephanie Chen, Yannis Kalantidis, Li Jia-Li, David Ayman Shamma, Michael Bernstrein, Li Fei-Fei)

- [Visual Hull Data Sets](#)a collection of visual hull datasets (Svetlana Lazebnik, Yasutaka Furukawa, and Jean Ponce)

14、People (static), human body pose

- [Frames Labeled In Cinema (FLIC)](#) - 20928 frames labeled with human pose (Sapp, Taskar)
- [Leeds Sports Pose Dataset](#) - 2000 pose annotated images of mostly sports people (Johnson, Everingham)
- [MPII Human Pose Dataset](#) - 25K images containing over 40K people with annotated body joints, 410 human activities {Andriluka, Pishchulin, Gehler, Schiele)
- [MPII Human Pose Dataset](#) - MPII Human Pose dataset is a de-facto standard benchmark for evaluation of articulated human pose estimation. (Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, Bernt Schiele)
- [People In Photo Albums](#) - Social media photo dataset with images from Flickr, and manual annotations on person heads and their identities.(Ning Zhang and Manohar Paluri and Yaniv Taigman and Rob Fergus and Lubomir Bourdev)
- [Person Recognition in Personal Photo Collections](#) - we introduced three harder splits for evaluation and long-term attribute annotations and per-photo timestamp metadata.(Oh, Seong Joon and Benenson, Rodrigo and Fritz, Mario and Schiele, Bernt)
- [Pointing'04 ICPR Workshop Head Pose Image Database](#)
- [UC-3D Motion Database](#) - Available data types encompass high resolution Motion Capture, acquired with MVN Suit from Xsens and Microsoft Kinect RGB and depth images.(Institute of Systems and Robotics, Coimbra, Portugal)
- [VGG Human Pose Estimation datasets](#) including the BBC Pose (20 videos with an overlaid sign language interpreter), Extended BBC Pose (72 additional training videos), Short BBC Pose (5 one hour videos with sign language signers), and ChaLearn Pose (23 hours of Kinect data of 27

persons performing 20 Italian gestures). (Charles, Everingham, Pfister, Magee, Hogg, Simonyan, Zisserman)

15、People Detection and Tracking Databases

- 3D KINECT Gender Walking data base (L. Igual, A. Lapedriza, R. Borràs from UB, CVC and UOC, Spain)
- AGORASET: a dataset for crowd video analysis (Nicolas Courty et al)
- CASIA gait database (Chinese Academy of Sciences)
- CAVIAR project video sequences with tracking and behavior ground truth (CAVIAR team/Edinburgh University - EC project IST-2001-37540)
- CMU Panoptic Studio Dataset - Multiple people social interaction dataset captured by 500+ synchronized video cameras, with 3D full body skeletons and calibration data. (H. Joo, T. Simon, Y. Sheikh)
- CUHK Crowd Dataset - 474 video clips from 215 crowded scenes (Shao, Loy, and Wang)
- CUHK01 Dataset : Person re-id dataset with 3, 884 images of 972 pedestrians (Rui Zhao et al)
- CUHK02 Dataset : Person re-id dataset with five camera view settings. (Rui Zhao et al)
- CUHK03 Dataset : Person re-id dataset with 13,164 images of 1,360 pedestrians (Rui Zhao et al)
- Caltech Pedestrian Dataset (P. Dollar, C. Wojek, B. Schiele and P. Perona)
- Daimler Pedestrian Detection Benchmark 21790 images with 56492 pedestrians plus empty scenes (D. M. Gavrila et al)
- Driver Monitoring Video Dataset (RobeSafe + Jesus Nuevo-Chiquero)
- DukeMTMC: Duke Multi-Target Multi-Camera tracking dataset - 8 cameras, 85 min, 2m frames, 2000 people of video (Ergys Ristani, Francesco Solera, Roger S. Zou, Rita Cucchiara, Carlo Tomasi)
- Edinburgh overhead camera person tracking dataset (Bob Fisher, Bashia Majecka, Gurkirt Singh, Rowland Sillito)
- GVVPerfcapEva - repository of human shape and performance capture data, including full body skeletal, hand tracking, body shape, face performance, interactions (Christian Theobalt)
- HAT database of 27 human attributes (Gaurav Sharma, Frederic Jurie)

- INRIA Person Dataset (Navneet Dalal)
- Inria Dressed human bodies in motion benchmark - Benchmark containing 3D motion sequences of different subjects, motions, and clothing styles that allows to quantitatively measure the accuracy of body shape estimates.(Jinlong Yang, Jean-Sbastien Franco, Franck H=E9troy-Wheeler, and Stefanie Wuhrer)
- Izmir - omnidirectional and panoramic image dataset (with annotations) to be used for human and car detection (Yalin Bastanlar)
- Joint Attention in Autonomous Driving (JAAD) - The dataset includes instances of pedestrians and cars intended primarily for the purpose of behavioural studies and detection in the context of autonomous driving.(Iuliia Kotseruba, Amir Rasouli and John K. Tsotsos)
- MAHNOB: MHI-Mimicry database - A 2 person, multiple camera and microphone database for studying mimicry in human-human interaction scenarios. (Sun, Lichtenauer, Valstar, Nijholt, and Pantic)
- MIT CBCL Pedestrian Data (Center for Biological and Computational Learning)
- MPI DYNA - A Model of Dynamic Human Shape in Motion (Max Planck Tubingen)
- MPI FAUST Dataset A data set containing 300 real, high-resolution human scans, with automatically computed ground-truth correspondences (Max Planck Tubingen)
- MPI MOSH Motion and Shape Capture from Markers. MOCAP data, 3D shape meshes, 3D high resolution scans. (Max Planck Tubingen)
- Market-1501 Dataset - 32,668 annotated bounding boxes of 1,501 identities from up to 6 cameras (Liang Zheng et al)
- Modena and Reggio Emilia first person head motion videos (Univ of Modena and Reggio Emilia)
- Multimodal Activities of Daily Living - including video, audio, physiological, sleep, motion and plug sensors. (Alexia Briasouli)
- Multiple Object Tracking Benchmark - A collection of datasets with ground truth, plus a performance league table (ETHZ, U. Adelaide, TU Darmstadt)
- Multispectral visible-NIR video sequences - Annotated multispectral video, visible + NIR (LE2I, Universit de Bourgogne)
- NYU Multiple Object Tracking Benchmark (Konrad Schindler et al)

- [Occluded Articulated Human Body Dataset](#) - Body pose extraction and tracking under occlusions, 6 RGB-D sequences in total (3500 frames) with one, two and three users, marker-based ground truth data(Markos Sigalas, Maria Pateraki, Panos Trahanias)
- [PARSE Dataset Additional Data](#) - facial expression, gaze direction, and gender (Antol, Zitnick, Parikh)
- [PARSE Dataset of Articulated Bodies](#) - 300 images of humans and horses (Ramanan)
- [PETS 2009 Crowd Challange dataset](#) (Reading University & James Ferryman)
- [PETS Winter 2009 workshop data](#) (Reading University & James Ferryman)
- [PETS: Performance Evaluation of Tracking and Surveillance](#) (Reading University & James Ferryman)
- [PIROPO](#) - People in Indoor ROoms with Perspective and Omnidirectional cameras, with more than 100,000 annotated frames (GTI-UPM, Spain)
- [People-Art](#) - a databased containing people labelled in photos and artwork (Qi Wu and Hongping Cai)
- [Photo-Art-50](#) - a databased containing 50 object classes annotated in photos and artwork (Qi Wu and Hongping Cai)
- [Pixel-based change detection benchmark dataset](#) (Goyette et al)
- [RAiD](#) - Re-Identification Across Indoor-Outdoor Dataset: 43 people, 4 cameras, 6920 images (Abir Das et al)
- [SYNTHIA](#) - Large set (~half million) of virtual-world images for training autonomous cars to see. (ADAS Group at Computer Vision Center)
- [Shinpuhkan 2014](#) - A Person Re-identification dataset containing 22,000 images of 24 people captured by 16 cameras. (Yasutomo Kawanishi et al.)
- [Stanford Structured Group Discovery dataset](#) - Discovering Groups of People in Images (W. Choi et al)
- [Temple Color 128 - Color Tracking Benchmark](#) - Encoding Color Information for Visual Tracking (P. Liang, E. Blasch, H. Ling)
- [Transient Biometrics Nails Dataset V01](#) (Igor Barros Barbosa)
- [Univ of Central Florida - Crowd Dataset](#) (Saad Ali)
- [Univ of Central Florida - Crowd Flow Segmentation datasets](#) (Saad Ali)

- [VIPeR: Viewpoint Invariant Pedestrian Recognition](#) - 632 pedestrian image pairs taken from arbitrary viewpoints under varying illumination conditions. (Gray, Brennan, and Tao)
- [Visual object tracking challenge datasets](#) - The VOT datasets is a collection of fully annotated visual object tracking datasets used in the single-target short-term visual object tracking challenges.(The VOT committee)
- [WIDER Attribute Dataset](#) - WIDER Attribute is a large-scale human attribute dataset, with 13789 images belonging to 30 scene categories, and 57524 human bounding boxes each annotated with 14 binary attributes.(Li, Yining and Huang, Chen and Loy, Chen Change and Tang, Xiaoou)

16、Remote Sensing

- [Brazilian Cerrado-Savanna Scenes Dataset](#) - Composition of IR-R-G scenes taken by RapidEye sensor for vegetation classification in Brazilian Cerrado-Savanna. (K. Nogueira, J. A. dos Santos, T. Fornazari, T. S. Freire, L. P. Morellato, R. da S. Torres)
- [Brazilian Coffee Scenes Dataset](#) - Composition of IR-R-G scenes taken by SPOT sensor for identification of coffee crops in Brazilian mountains.(O. A. B. Penatti, K. Nogueira, J. A. dos Santos.)
- [Building Detection Benchmark](#) -14 images acquired from IKONOS (1 m) and QuickBird (60 cm)(Ali Ozgun Ok and Caglar Senaras)
- [CBERS-2B, Landsat 5 TM, Geoeye, Ikonos-2 MS and ALOS-PALSAR](#) - land-cover classification using optical images(D. Osaku et al.)
- [Furnas and Tiete](#) - sediment yield classification(Pisani et al.)
- [ISPRS 2D semantic labeling](#) - Height models and true ortho-images with a ground sampling distance of 5cm have been prepared over the city of Potsdam/Germany (Franz Rottensteiner, Gunho Sohn, Markus Gerke, Jan D. Wegner)
- [ISPRS 3D semantic labeling](#) - nine class airborne laser scanning data (Franz Rottensteiner, Gunho Sohn, Markus Gerke, Jan D. Wegner)
- [The Linkoping Thermal InfraRed dataset](#) - The LTIR dataset is a thermal infrared dataset for evaluation of Short-Term Single-Object (STSO) tracking (Linkoping University)

17、Scene Segmentation or Classification

- [Barcelona](#) - 15,150 images, urban views of Barcelona (Tighe and Lazebnik)
- [CMU Visual Localization Data Set](#) - Dataset collected over the period of a year using the Navlab 11 equipped with IMU, GPS, INS, Lidars and cameras.(Hernan Badino, Daniel Huber and Takeo Kanade)
- [COLD (COsy Localization Database) - place localization](#) (Ullah, Pronobis, Caputo, Luo, and Jensfelt)
- [EDUB-Seg](#)- Egocentric dataset for event segmentation.(Mariella Dimiccoli, Marc Bolaños, Estefania Talavera, Maedeh Aghaei, Stavri G. Nikolov, and Petia Radeva.)
- [Fifteen Scene Categories](#)a dataset of fifteen natural scene categories (Fei-Fei Li and Aude Oliva)
- [Geometric Context - scene interpretation images](#) (Derek Hoiem)
- [Indoor Place Recognition Dataset for localization of Mobile Robots](#) - The dataset contains 17 different places built from 2 different robots (virtualMe and pioneer) (Raghavender Sahdev, John K. Tsotsos.)
- [Indoor Scene Recognition](#) - 67 Indoor categories, 15620 images (Quattoni and Torralba)
- [LM+SUN](#) - 45,676 images, mainly urban or human related scenes (Tighe and Lazebnik)
- [Maritime Imagery in the Visible and Infrared Spectrums](#) - VAIS contains simultaneously acquired unregistered thermal and visible images of ships acquired from piers (Zhang, M.M, Choi, J., Daniilidis, K., Wolf, M.T. & Kanan)
- [NYU V2 Mixture of Manhattan Frames Dataset](#) - We provide the Mixture of Manhattan Frames (MMF) segmentation and MF rotations on the full NYU depth dataset V2 by Silberman et al. (Straub, Julian and Rosman, Guy and Freifeld, Oren and Leonard, John J. and Fisher III, John W.)
- [Places Scene Recognition database](#) - 205 scene categories and 2.5 millions of images (Zhou, Lapedriza, Xiao, Torralba, and Oliva)
- [RGB-NIR Scene Dataset](#) - 477 images in 9 categories captured in RGB and Near-infrared (NIR) (Brown and Susstrunk)
- [SUN 2012](#) - 16,873 fully annotated scene images for scene categorization (Xiao et al)
- [SUN 397](#) - 397 scene categories for scene classification (Xiao et al)

- [SUN RGB-D: A RGB-D Scene Understanding Benchmark Suite](#) - 10,000 RGB-D images, 146,617 2D polygons and 58,657 3D bounding boxes (Song, Lichtenberg, and Xiao)
- [SYNTHIA](#) - Large set (~half million) of virtual-world images for training autonomous cars to see. (ADAS Group at Computer Vision Center)
- [Sift Flow (also known as LabelMe Outdoor, LMO)](#) - 2688 images, mainly outdoor natural and urban (Tighe and Lazebnik)
- [Stanford Background Dataset](#) - 715 images of outdoor scenes containing at least one foreground object (Gould et al)
- [Surface detection](#) - Real-time traversable surface detection by colour space fusion and temporal analysis - Evaluation Dataset (Breckon, Toby P., Katramados, Ioannis)
- [UIUC Fifteen Scene Categories](#) – Dataset of 15 natural scene categories.
- [ViDRILO](#) - ViDRILO is a dataset containing 5 sequences of annotated RGB-D images acquired with a mobile robot in two office buildings under challenging lighting conditions.(Miguel Cazorla ,J. Martinez-Gomez, M. Cazorla, I. Garcia-Varea and V. Morell.)
- [Video Object Segmentation dataset DAVIS](#) - Densely Annotated VIdeo Segmentation (F. Perazzi, J. Pont-Tuset, B. McWilliams, L. Van Gool, M. Gross, and A. Sorkine-Hornung)

18、Segmentation(General)

- [Alpert et al. Segmentation evaluation database](#) (Sharon Alpert, Meirav Galun, Ronen Basri, Achi Brandt)
- [BMC (Background Model Challenge)](#) - A dataset for comparing background subtraction algorithms, comp= osed of real and synthetic videos(Antoine)
- [Berkeley Segmentation Dataset and Benchmark](#) (David Martin and Charless Fowlkes)
- [CTU Color and Depth Image Dataset of Spread Garments](#) - Images of spread garments with annotated corners.(Wagner, L., Krejov D., and Smutn V. (Czech Technical University in Prague))
- [CTU Garment Folding Photo Dataset](#) - Color and depth images from various stages of garment folding.(Sushkov R., Melkumov I., Smutn y V. (Czech Technical University in Prague))

- [DeformIt 2.0](#) - Image Data Augmentation Tool: Simulate novel images with ground truth segmentations from a single image-segmentation pair (Brian Booth and Ghassan Hamarneh)
- [GrabCut Image database](#) (C. Rother, V. Kolmogorov, A. Blake, M. Brown)
- [LabelMe images database and online annotation tool](#) (Bryan Russell, Antonio Torralba, Kevin Murphy, William Freeman)
- [PetroSurf3D](#) - 26 high resolution (sub-millimeter accuracy) 3D scans of rock art with pixelwise labeling of petroglyphs for segmentation(Poier, Seidl, Zeppelzauer, Reinbacher, Schaich, Bellandi, Marretta, Bischof)
- [SYNTHIA](#) - Large set (~half million) of virtual-world images for training autonomous cars to see. (ADAS Group at Computer Vision Center)
- [Stony Brook University Shadow Dataset (SBU-Shadow5k)](#) - Large scale shadow detection dataset from a wide variety of scenes and photo types, with human annotations (Tomas F.Y. Vicente, Le Hou, Chen-Ping Yu, Minh Hoai, Dimitris Samaras)

19、Simultaneous Localization and Mapping

- [Event-based Data for Pose Estimation, Visual Odometry, and SLAM](#)The data also include intensity images, inertial measurements, and ground truth from a motion-capture system. (ETH)
- [RAWSEEDS SLAM benchmark datasets](#) (Rawseeds Project)
- [TUM RGB-D Benchmark](#) - Dataset and benchmark for the evaluation of RGB-D visual odometry and SLAM algorithms (BCrgen Sturm, Nikolas Engelhard, Felix Endres, Wolfram Burgard and Daniel Cremers)
- [Visual Odometry / SLAM Evaluation](#) - The odometry benchmark consists of 22 stereo sequences (Andreas Geiger and Philip Lenz and Raquel Urtasun)

20、Surveillance

- [AVSS07: Advanced Video and Signal based Surveillance 2007 datasets](#) (Andrea Cavallaro)
- [Activity modeling and abnormality detection dataset](#) - The dataset containes a 45 minutes video with annotated anomalies.(Jagan Varadarajan and Jean-Marc Odobez)

- [Background subtraction](#) - a list of datasets about background subtraction(Thierry BOUWMANS)
- [CMUSRD: Surveillance Research Dataset](#) - multi-camera video for indoor surveillance scenario (K. Hattori, H. Hattori, et al)
- [DukeMTMC: Duke Multi-Target Multi-Camera tracking dataset](#) - 8 cameras, 85 min, 2m frames, 2000 people of video (Ergys Ristani, Francesco Solera, Roger S. Zou, Rita Cucchiara, Carlo Tomasi)
- [ETISEO Video Surveillance Download Datasets](#) (INRIA Orion Team and others)
- [MAHNOB Databases](#) -including Laughter Database,HCI-tagging Database,MHI-Mimicry Database(M. Pantic. etc.)
- [Multispectral visible-NIR video sequences](#) - Annotated multispectral video, visible + NIR (LE2I, Universit de Bourgogne)
- [Openvisor - Video surveillance Online Repository](#) (Univ of Modena and Reggio Emilia)
- [Parking-Lot dataset](#) - Parking-Lot dataset is a car dataset which focus on moderate and heavily occlusions on cars in the parking lot scenario.(B. Li, T.F. Wu and S.C. Zhu)
- [Pornography Database](#) - The Pornography database is a pornography detection dataset containing nearly 80 hours of 400 pornographic and 400 non-pornographic videos extracted from pornography websites and Youtube. (Sandra Avila, Eduardo Valle, Arnaldo de A.)
- [Queen Mary Multi-Camera Distributed Traffic Scenes Dataset (QMDTS)](#) - The QMDTS is collected from urban surveillance environment for the study of surveillance behaviours in distributed scenes.(Dr. Xun Xu. Prof. Shaogang Gong and Dr. Timothy Hospedales)
- [SALSA: Synergetic sociAL Scene Analysis](#) - A Novel Dataset for Multimodal Group Behavior Analysis(Xavier Alameda-Pineda etc.)
- [SBMnet (Scene Background Modeling.NET)](#) - A dataset for testing background estimation algorithms(Pierre-Marc Jodoin, Lucia Maddalena, and Alfredo Petrosino)
- [SCOUTER](#) - video surveillance ground truthing (shifting perspectives, different setups/lighting conditions, large variations of subject). 30 videos and approximately 36,000 manually labeled frames. (Catalin Mitrea)

- [SJTU-BEST](#)One surveillance-specified datasets platform with realistic, on-using camera-captured, diverse set of surveillance images and videos (Shanghai Jiao Tong University)
- [SPEVI: Surveillance Performance EValuation Initiative](#) (Queen Mary University London)
- [Shinpuhkan 2014](#) - A Person Re-identification dataset containing 22,000 images of 24 people captured by 16 cameras. (Yasutomo Kawanishi et al.)
- [Tracking in extremely cluttered scenes](#) - this single object tracking dataset has 28 highly cluttered sequences with per frame annotation(Jingjing Xiao,Linbo Qiao,Rustam Stolkin,Ale Leonardis)
- [UCSD Anomaly Detection Dataset](#) - a stationary camera mounted at an elevation, overlooking pedestrian walkways, with unusual pedestrian or non-pedestrian motion.
- [UCSD trajectory clustering and analysis datasets](#) - (Morris and Trivedi)
- [Udine Trajectory-based anomalous event detection dataset](#) - synthetic trajectory datasets with outliers (Univ of Udine Artificial Vision and Real Time Systems Laboratory)
- [WIDER Attribute Dataset](#) - WIDER Attribute is a large-scale human attribute dataset, with 13789 images belonging to 30 scene categories, and 57524 human bounding boxes each annotated with 14 binary attributes.(Li, Yining and Huang, Chen and Loy, Chen Change and Tang, Xiaoou)

21、Textures

- [Brodatz Texture, Normalized Brodatz Texture, Colored Brodatz Texture, Multiband Brodatz Texture](#) 154 new images plus 112 original images with various transformations (A. Safia, D. He)
- [Color texture images by category](#) (textures.forrest.cz)
- [Columbia-Utrecht Reflectance and Texture Database](#) (Columbia & Utrecht Universities)
- [DynTex: Dynamic texture database](#) (Renaud Piteri, Mark Huiskes and Sandor Fazekas)
- [KTH TIPS & TIPS2 textures](#) - pose/lighting/scale variations (Eric Hayman)
- [Oulu Texture Database](#) (Oulu University)

- Oxford Describable Textures Dataset - 5640 images in 47 categories (M.Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, A. Vedaldi)
- Prague Texture Segmentation Data Generator and Benchmark (Mikes, Haindl)
- Texture DatabaseThe texture database features 25 texture classes, 40 samples each(Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce)
- Uppsala texture dataset of surfaces and materials - fabrics, grains, etc.
- Vision Texture (MIT Media Lab)

22、Urban Datasets

- Barcelona - 15,150 images, urban views of Barcelona (Tighe and Lazebnik)
- CMP Facade Database - Includes 606 rectified images of facades from various places with 12 architectural classes annotated.(Radim Tylecek)
- LM+SUN - 45,676 images, mainly urban or human related scenes (Tighe and Lazebnik)
- MIT CBCL StreetScenes Challenge Framework: (Stan Bileschi)

23、Handwritten Digits

- MNIST – large dataset containing a training set of 60,000 examples, and a test set of 10,000 examplespedes

24、Feature Detection and Description

- VGG Affine Dataset – Widely used dataset for measuring performance of feature detection and description. CheckVLBenchmarks for an evaluation framework.Benchmarking Activity Recognition – CVPR 2012 tutorial covering various datasets for action recognition.