1. Image pyramids (scale-space)

- Reading material
  - Szeliski Chapter 4.1 (download book here: https://www.cs.ccu.edu.tw/~damon/tmp/SzeliskiBook_20100903_draft.pdf)
  - "Pyramid methods in image processing" by Adelson et al (download here: http://persci.mit.edu/pub_pdfs/RCA84.pdf)
  - "Distinctive Image Features from Scale-Invariant Keypoints" by David Lowe (download here: http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf)
- Supplemental reading
  - "Feature Detection with Automatic Scale Selection" by T. Lindeberg (download here: ftp://ftp.nada.kth.se/CVAP/reports/cvap198.pdf)
  - "Pyramid methods in image processing" by Adelson et al (download here: http://persci.mit.edu/pub_pdfs/RCA84.pdf)
- Slides
  - Feature detection and matching
- Exercises
  - Lab2
- Matlab code (see folder Demos & Videos on Blackboard)
  - Harris corner detector
  - Gaussian scale space
  - Difference of Gaussian (DoG) scale space
- Learning objectives
  - Mention typical applications of feature detection, description, and matching.
  - Explain what an "interest point" is, and what such points can be used for.
  - Feature detection
    - Explain how the Harris corner detector works (at least for an axis-aligned corner…).
    - Explain how the blob detector works (i.e., the feature detection strategy used in SIFT).
  - Motivate the importance of "scale invariance" and "rotational invariance".
  - Describe strategies for obtaining scale invariance and rotational invariance (e.g., automatic scale selection).
  - Describe different strategies for "feature description" (e.g., image patches and the SIFT descriptor).
  - Describe the basic principle of "feature matching".
  - Mention typical applications of image pyramids.
  - Explain the steps involved in constructing a Gaussian pyramid.
  - Explain the link between image scale and the standard deviation of the Gaussian filter.
  - Explain how to construct a difference of Gaussian (DoG) pyramid.
  - Argue that the difference of Gaussian (DoG) operator is a reasonable approximation to the Laplacian of Gaussian (LoG) operator.
  - Describe the typical usage of DoG pyramids in feature detection (what type of feature are we typically looking for? How?).
  - Relate scale space to
    - feature detection, description, and matching
    - template matching
    - motion estimation

2. Image formation

- Reading material
  - Szeliski Chapter 2, pages 31-40 + 46-60
  - "Projective geometry for machine vision" by Mundy & Zisserman, pages 468-483 + 507-511 + section 23.10.6 (download here: http://www.cs.cmu.edu/~ph/869/papers/zisser-mundy.pdf)
- Supplemental reading:
  - "Flexible Camera Calibration By Viewing a Plane From Unknown Orientations" by Zhengyou Zhang (download here: http://www.vision.caltech.edu/bouguetj/calib_doc/papers/zhan99.pdf)
- Slides
  - Image formation
- Exercises
  - Lab3
- Learning objectives
  - Cameras
    - Describe the geometry of the camera model/pinhole camera (e.g., center of projection, projection plane, principal point, focal length, etc.).
    - Explain the role of the calibration matrix (K) and the camera intrinsics.
    - Explain the role of the camera matrix (P) and the camera extrinsics.
    - Mention at least one strategy for camera calibration.
  - Projective geometry
    - Explain the geometry under perspective viewing (e.g., view point, rays, vanishing points, and horizon).
    - Mention examples of properties that remain – and do not remain – invariant under perspective viewing.
    - Explain and motivate the use of homogenous coordinates in projective geometry.
    - Define what is meant by a homography (=projective transformation), and know that it is only defined up to a scale.
    - Explain the basic steps involved in solving for an unknown homography using point correspondences.
    - Explain the basic steps in image rectification.
    - Argue that panoramic views can be generated using homographies
      - when the camera centers are fixed (always!)
      - when the camera centers change AND the imaged world object is planar.
    - Explain how to constrain feature matching using a global motion model (e.g., a homography) and RANSAC

3. Stereo and epipolar geometry

- Reading material
  - Ch. 7.1-7.2 (p. 345-348, 353-354), Ch. 11.1-11.3 (537-540, 543-548)
  - "Multiple View Geometry in Computer Vision (Second Edition)" by Hartley and Zisserman: Ch. 9.1-9.3 (download here: http://www.robots.ox.ac.uk/~vgg/hzbook/hzbook2/HZepipolar.pdf)
  - "Projective geometry for machine vision" by Mundy & Zisserman, pages 468-483 + 507-511 + section 23.10.6 (download here: http://www.cs.cmu.edu/~ph/869/papers/zisser-mundy.pdf)
- Slides
  - Stereo-Epipolar Geometry and Image Alignment
- Exercises
  - Lab4
- Learning objectives
  - Explain the term disparity, for instance using the parallel two-camera setup.
  - Know what is meant by triangulation.
  - Explain the epipolar geometry for two camera views (e.g., epipoles, epipolar lines, epipolar plane).
  - Describe the epipolar constraint for calibrated and uncalibrated cameras, respectively (i.e., essential matrix versus fundamental matrix).
  - Motivate the use of stereo image rectification.
  - Describe the process followed in triangulation. How many points are needed in order to calculate the real-world coordinates (X)?
  - Explain how to compute the fundamental matrix (i.e., stereo pipeline with weak calibration).
- Structure from motion (SFM)
  - Define the SFM problem.
  - Explain what happens to the reconstruction if there are no constraints on the camera matrix or on the scene.
  - Describe how to solve the SFM problem with an affine camera and under orthography.
  - Explain how to correct for affine ambiguity.
- Photometric stereo
  - Explain the basic idea underlying photometric stereo (or “shape from light”).
  - Describe the basic principle of “depth from normals”.
- Shading
  - Describe the basics of light transport
  - Illumination -> reflectance -> perception
  - Light sources
  - Types of reflection
  - Describe how shape from shading works
  - Mention strategies for estimating light source directions.

4. Edges and lines

- Reading material
  - Szeliski chapter 4.2.-4.3
  - "Learning to detect natural image boundaries using local brightness, color, and texture cues" by Martin et al. (download here: http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/papers/mfm-pami-boundary.pdf)
  - Holistically-Nested Edge Detection", S. Xie and Z. Tu (download from: https://arxiv.org/pdf/1504.06375.pdf)
- Slides
  - Edges and lines
- Exercises
  - Lab5
- Learning objectives
  - Explain how to detect edges using a gradient operator.
  - Describe how to detect edges using the Laplacian of Gaussian (LoG) operator.
  - Motivate the importance of prefiltering the image before gradient-based edge detection.
  - Explain how the Canny operator works with emphasis on non-maximum suppression and hysteresis.
  - Explain how to generate a texton map.
  - Explain how to detect edges between two textures using a half-disc detector.
  - List different strategies for linking edges (for example successive approximation of lines/curves) and edge thinning.
  - Describe how supervised edge detectors work.
  - Describe how Hough Transform works and on what problems it can be applied to.

5. Segmentation

- Reading material
  - Szeliski chapter 5.1.1, 5.1.4, 5.2-5.5
  - "Gradient Vector Flow: A New External Force for Snakes", Xu and Prince, until page 68, focuse on energy Eq. 1, the equivalent force balance Eq. 6, and the GVF field Eq. 10 (download from: http://iacl.ece.jhu.edu/pubs/p087c.pdf)
- Slides
  - Segmentation
- Exercises
  - Lab6
- Learning objectives
  - Motive the use of clustering for image segmentation.
  - Mention typical image features used for clustering.
  - Explain the difference between divisive clustering and agglomerative clustering.
  - Describe the K-means algorithm.
  - Explain how the Mean Shift algorithm works.
  - For graph-based segmentation you should know the following:
    - An image can be represented as a graph, where the pixels are the nodes and edges connect pixels.
    - Typically, only neighboring pixels are connected.
    - In our context, graphs can be thought of as “roadmaps”, where the nodes represent cities and the edges represent roads between cities. Each edge is associated with a certain cost, which in the roadmap example could be the distance (in km) between two connected cities.
    - Normalized cuts: For a given path from A to B (i.e., a connected series of edges leading from node A to node B), we may encounter some “obstacle” that makes the total cost go very high. This is the basic idea underlying the normalized cut segmentation algorithm; the obstacles could be edges in the image. In this case, the normalized cut algorithm would group together pixels that are similar in appearance (e.g., color) and surrounded by an edge.
    - Shortest path: We may also define the cost of the edges such that the cost of the edge connecting two neighboring pixels is low if both pixels lie on an edge (and high otherwise). If both A and B lie on the same edge, we can then find the edge connecting A and B by finding the shortest path (using Dijsktra’s algorithm) from A to B. This is the basic idea underlying “intelligent scissors” or “livewire”.
  - Motivate the use of snakes and level sets for image segmentation.
  - Explain the basic concept of snakes, including physical interpretation of the energy terms.
  - Given an initial contour, describe in words how the final contour (snake) is obtained.
  - Describe the Gradient Vector Flow (GVF) field and illustrate the advantages of this field over the traditional external force field.
  - Describe the main differences between snakes and level sets, including the contour representation used in the two methods.
  - Describe how the level set is moved in each iteration (i.e., equation of motion).

6. Clustering and Subspace Learning

- Reading material
  - "Elements of Statistical Learning (2nd ed.)", Chapters 14.3.4 - 14.5.3 (pages 507-547), 14.6 - 14.7 (pages 553-570), 4.3 (pages 106-119) (download book here: https://web.stanford.edu/~hastie/Papers/ESLII.pdf)
  - "Understanding Machine Learning: From Theory to Algorithms", Chapters 23.2 - 23.3 (pages 329-331) (download here: http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/copy.html)
  - "The EM Algorithm for Gaussian Mixtures" (download from here: http://www.ics.uci.edu/~smyth/courses/cs274/notes/EMnotes.pdf)
- Supplemental reading:
  - Graph Embedding and Extensions", (paper: http://ieeexplore.ieee.org/document/4016549/)
  - "Generalized Multi-view Embedding", (paper https://arxiv.org/pdf/1605.09696.pdf)
  - "Class-specific Reference Discriminant Analysis" (download here: http://ieeexplore.ieee.org/document/6998872/)
- Slides
  - Clustering and Subspace Learning
- Exercises
  - Lab7
- Learning objectives
  - Describe the difference of in-sample and out-of-sample analysis.
  - Describe what is model selection, optimization and evaluation.
  - Describe the hierarchical process followed to select a ML model
    - Problem type identification
    - Type of solutions applied (linear vs nonlinear)
    - Other constraints
    - Type of models
    - Data collection and parameters optimization
  - Describe the difference between convex vs non-convex optimization.
  - Describe K-Means clustering (model's parameters, and how they are optimized)
  - Describe Mixture of Gaussians (model's parameters, and how they are optimized)
  - What is K-Medoids
  - Describe in words the two hierarchical clustering approaches
  - What is Self-Organizing Map?
  - Describe the problem of Dimensionality Reduction
  - Principal Component Analysis (model's parameters, how they are optimized)
  - Non-negative Matrix Factorization (model's parameters, how they are optimized)
  - Independent Component Analysis (model's parameters, differences from PCA and NMF, generic description of the optimization process)
  - What are Random Projections?
  - Linear Discriminant Analysis (model's parameters, how they are optimized)
  - Graph-based subspace learning:
    - Describe the generic idea of using an intrinsic and a penalty graph for subspace learning
    - Define the graph G = {V,E}; how is the data used to form a graph?
    - Affinity and Weight matrix of a graph
    - What is the Laplacian Matrix?
    - Express LDA as a graph-based subspace learning method
  - Describe Spectral Clustering method; What is it's difference from K-Means?
  - Describe in words what is Multi-view Embedding and why it is used.

7. Regression, Classification and non-linear models

- Reading material
  - "Elements of Statistical Learning (2nd ed.)", Chapters 3.2, 3.4.1, 3.4.2, 3.5.1, 4.2-4.4, 12.1-12.5
  - "Using the Nystrom method to Speed Up Kernel Machines", (download from http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=028D072DC62B6EB05AF99C94FEB85688?doi=10.1.1.18.7519&rep=rep1&type=pdf)
- Supplemental reading:
  - "Nonlinear Projection Trick in kernel methods", (download from http://mipal.snu.ac.kr/images/4/47/NPT.pdf)
  - "Random Features for Large-Scale Kernel Machines", (download from https://people.eecs.berkeley.edu/~brecht/papers/07.rah.rec.nips.pdf)
  - "Graph Embedding and Extensions", (download from http://ieeexplore.ieee.org/document/4016549/)
  - "A Survey of Randomized Algorithms for Training Neural Networks", (download from https://www.sciencedirect.com/science/article/pii/S002002551600058X)
- Slides
  - Regression and Classification
- Exercises
  - Lab8
- Learning objectives
  - Describe the Linear Regression model (model's parameters, and how they are optimized).
  - Describe the Ridge Regression model (difference from linear regression, model's parameters, how they are optimized).
  - What is LASSO? Connection and differences from the above models.
  - Principal Components Regression
  - Describe how can a regression model be used for classification.
  - Describe the Logistic Regression model (model's parameters, and how they are optimized); Describe why it is a probability-like model.
  - Describe the two ways that can be used to extend linear models to non-linear ones
    - Use of explicit nonlinear mappings
    - Use of implicit nonlinear mappings (kernels)
  - Describe RBF networks and its randomized counterparts.
  - Describe the Support Vector Machine classifier.
  - Describe what is a kernel function and what are the properties that a matrix needs to satisfy in order to be a kernel matrix.
  - Describe Kernel Discriminant Analysis as an extension of Linear Discriminant Analysis (using the Representer Theorem $\mathbf{W} = \mathbf{\Phi A}$ and the definition of the kernel matrix $\mathbf{K} = \mathbf{\Phi}^T\mathbf{\Phi}$)
  - Describe how we can calculate the Eulidean distance between two vectors and how this can be used to extend k-Means to its kernel-based version.
  - Describe Kernel-based Regression.
  - What is the Nonlinear Projection Trick?
  - What is Nyström Approximation? In which cases do we use it?

8. Neural Networks

- Reading material
  - "Understanding Machine Learning: From Theory to Algorithms", Chapters 11, 14, 20 (download here: http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/copy.html)
- Supplemental reading:
  - "Convolutional Neural Networks for Visual Recognition", Fei-Fei Li, (http://cs231n.github.io/)
    - Backpropagation, Intuitions
    - Neural Networks Part 2 and 3
    - Convolutional Neural Networks, Convolution / Pooling Layers)
- Slides
  - Neural Networks
- Exercises
  - Lab9
- Learning objectives
  - Describe the building blocks of a feedforward neural network.
  - Describe what is a Single-hidden Layer Feedforward Neural network (model's parameters and how the output is calculated).
  - Describe the importance of the activation function for neural networks.
  - Describe in words how the network's parameters are optimized (use the means square error criterion for the description)
  - Describe how the selection of the learning rate value can affect the optimization process of a network.
  - What is forward pass and what is backward pass in the Backpropagation algorithm? What signals are used in each of them and what functions?
  - Describe various data pre-processing steps used in training neural networks and why there is a need to pre-process the data.
  - What is learning with momentum?
  - What is learning with adaptive learning rate?
  - What is DropOut?
  - Describe what is the softmax function and how it is used in neural networks.
  - Describe what is the cross-entropy loss function.
  - Describe the building blocks of a Convolutional Neural Network.
  - Describe the parameters of a CNN
  - What is a convolutional neuron and what are its differences from a standard neuron?
  - Describe the training protocols
    - Train-validation-test split
    - Cross-validation and its variants

9. Scene Classification

- Reading material
    - “Multiple Instance Classification: review, taxonomy and comparative study”, Artificial Intelligence, 2013 (download from here: https://www.sciencedirect.com/science/article/pii/S0004370213000581)
    - “Support Vector Machines for Multiple-Instance Learning”, Neural Information Processing Systems, 2002 (download from here: https://proceedings.neurips.cc/paper/2002/file/3e6260b81898beacda3d16db379ed329-Paper.pdf)
    -
- Supplemental reading:
    - “A framework for multiple-instance learning”, Neural Information Processing Systems, 1998 (download from here: http://people.csail.mit.edu/tlp/publications/maron98framework.pdf)
    - “Multiple Instance Regression”, International Conference on Machine Learning, 2001 (download from here: http://engr.case.edu/ray_soumya/papers/mip.reg.icml01.pdf)
- Slides
    - Multiple Instance Learning
- Exercises
    - Lab10
- Learning objectives
    - Describe the problem if image classification and why (and how) it is modeled as a Multiple-Instance Learning (MIL) problem.
    - Explain the categorization of MIL methods in instance-level and bag-level methods.
    - Describe the standard assumption of instance-level MIL methods related to the distribution of image descriptors from different classes.
    - Describe how the instance-level MIL methods apply image classification.
    - Describe how bag-level MIL methods combine information from instances for image classification.
    - Describe the various distance/kernel functions used in MIL methods to exploit information from all instances.
    - Describe how embedding space classification works
    - What types of embedding are used in embedding-based classification?
    - Describe the Bag of Words (or Bag of Features) model
    - What is the difference between Bag of Words model and Discriminant Bag of Words models?
    - Describe a Discriminant Bag of Words (Features) model.

10. Object Recognition

- Reading material
  - Szeliski Ch. 14.1-14.3
  - "Rapid object detection using a boosted cascade of simple features" by Viola and Jones (download here: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.6807&rep=rep1&type=pdf).
  - "Histograms of Oriented Gradients for Human Detection" by Dalal and Triggs (download here: http://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf).
  - “Local Binary Patterns and Its Application to Facial Image Analysis: A Survey” (download from here: http://liris.cnrs.fr/Documents/Liris-5004.pdf)
  - "Pictorial structures for object recognition" by Felzenszwalb and Huttenlocher (download here: www.cs.cornell.edu/~dph/papers/pict-struct-ijcv.pdf)
  - “Selective Search for Object Recognition”, International Journal on Computer Vision, 2013 (download from here: http://www.huppelen.nl/publications/selectiveSearchDraft.pdf)
  - “Edge Boxes: Locating Object Proposals from Edges”, European Conference on Computer Vision, 2014 (download from here: https://web.bii.a-star.edu.sg/~zhangxw/files/EdgeBoxes_ECCV2014.pdf)
  - “Fast R-CNN”, arXiv:1504.08083, 2015 (download from here: https://arxiv.org/pdf/1504.08083.pdf)
  - “You Only Look Once: Unified, Real-Time Object Detection”, arXiv:1506.02640, 2015 (download from here: https://arxiv.org/abs/1506.02640)
- Supplemental reading:
  - "Active appearance models" by Cootes, Edwards, and Taylor (download here: https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/cootes-eccv-98.pdf)
  - “A survey on face detection in the wild: Past, present and future”, Computer Vision and Image Understanding, 2015 (download from here: https://www.sciencedirect.com/science/article/pii/S1077314215000727)
  - “Fast, Compact, and Discriminative: Evaluation of Binary Descriptors for Mobile Applications”, IEEE Transactions on Multimedia, 2017 (download here: http://ieeexplore.ieee.org/document/7583687/)
  - “Mask R-CNN”, arXiv:1703.06870, 2017 (download from here: https://arxiv.org/abs/1703.06870)
- Slides
  - Object Detection
- Exercises
  - Lab11
- Learning objectives:
  - Object detection
    - Know the difference between object detection and object identification
    - Describe the basic scheme used in object detection (such as face detection)
    - Describe the basic principle used when training a classifier for image classification (i.e. use both positive and negative examples)
    - Describe different types of features used for object detection (Viola-Jones, Histogram of Oriented Gradients, Local Binary Patterns)
    - Explain the basic principle of part-based models.
  - Active appearance models (can be used for object identification)
    - Mention applications of active appearance models
    - Explain what is meant by a shape model and a texture model, respectively
    - Describe the manual work involved in building/training an active appearance model.
    - Explain why we need to apply translation, rotation and scale correction (-Procrustes alignment) before applying PCA
  - Neural network-based object recognition
    - Describe a network architecture used for image classification
    - Describe the concept of receptive field of a convolutional filter of two different layers
    - What shapes do the convolutional filters take?
    - Describe the concept of Object Proposals methods
    - Describe the two approaches followed for object recognition/localization in images using CNNs (Region-based CNN and YOLO)

11. Tracking

- Reading material
  - Szeliski Ch 8.1 (except p. 397-398), 8.2 (except section 8.2.1).
  - “A Review of Visual Tracking”, Technical Report, 2008 (download from here: https://www.eecs.yorku.ca/research/techreports/2008/CSE-2008-07.pdf) , Ch. 2.1, 2.3, 5
  - “Pyramidal Implementation of Lucas Kanade Feature Tracker Description of the algorithm”, Intel Corporation, Microprocessor Research Labs, 2000 (download from here: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.185.585)
  - “Tracking-Learning-Detection”, IEEE Transactions on Pattern Analysis and Machine Learning, 2010 (download from here: http://ieeexplore.ieee.org/document/6104061/)
- Supplemental reading:
  - “Visual Tracking: An Experimental Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014 (download from here: http://ieeexplore.ieee.org/document/6671560/)
- Slides
  - Tracking
- Exercises
  - Lab12
- Learning objectives
  - Explain what Object Tracking in a video sequence is.
  - Describe different types of initialization for tracking
  - Describe what are the differences between trackers based on:
    - Point-wise predictions
    - Template predictions
    - Optical flow-based predictions
  - Describe the Kalman filter model and describe how it is used for tracking
  - List advantages and disadvantages of Kalman filter model for tracking
  - Describe Particle filter model and how it is used for tracking
  - Describe the Mean Shift Tracking model
  - Describe the Lucas-Kanade Tracker (KLT) model
  - Explain what the “brightness constancy assumption” is and how this leads to the optical flow equation $\nabla I \cdot u + It = 0$
  - Explain why we need to extend the KLT model to use multiple iterations
  - Describe the basic principles of Tracking-Learning-Detection (TLD) tracking model
  - Explain why the TLD model uses two experts (P and N). What are these experts?
  - Describe how the spatial and temporal structure is used from the P- and N-experts.

12. Salient Object Detection

- Reading material
    - "Salient Object Detection: A Survey" (download from here: https://arxiv.org/pdf/1411.5878.pdf)
    - "SLIC Superpixels Compared to State-of-the-art Superpixel Methods", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011 (download from here: https://infoscience.epfl.ch/record/177415/files/Superpixel_PAMI2011-2.pdf)
    - "Learning graph affinities for spectral graph-based salient object detection", Pattern Recognition, 2018 (download from here: https://www.sciencedirect.com/science/article/pii/S0031320316303570)
    - "Probabilistic Saliency Estimation", Pattern Recognition, 2018 (download from here: https://www.sciencedirect.com/science/article/pii/S0031320317303734)
    - "A Unified Approach to Salient Object Detection via Low Rank Matrix Recovery", Computer Vision and Pattern Recognition, 2012 (download from here: http://ieeexplore.ieee.org/document/6247758/)
    - "Fully Convolutional Networks for Semantic Segmentation", Computer Vision and Pattern Recognition, 2015 (download from here: http://ieeexplore.ieee.org/document/7298965/)
- Supplemental reading:
    - "Generic Promotion of Diffusion-based Salient Object Detection", International Conference on Computer Vision, 2015 (download from here: http://ieeexplore.ieee.org/document/7410390/?arnumber=7410390&tag=1)
    - "Automatic Object Segmentation by Quantum Cuts", International Conference on Pattern Recognition, 2012 (download from here: http://ieeexplore.ieee.org/document/6976740/)
- Slides
    - Salient Object Detection
- Exercises
    - Lab13
- Learning objectives
    - Explain what is the problem of Salient Object Detection (SOD), along with the different types of saliency problems.
    - What is the difference between intrinsic and extrinsic SOD methods?
    - Describe how the block-based segmentation methods work.
    - Describe how the region-based segmentation methods work.
    - Describe the saliency cues used by SOD methods.
    - Describe what is a superpixel.
    - Describe the Simple Linear Iterative Clustering (SLIC) method
    - Describe how the SOD problem can be modeled as a low-rank matrix problem.
    - Describe how an image can be modeled as a graph (which are the nodes and the edges, what do they express) for SOD
    - Describe the Diffusion-based SOD approach.
    - Describe the Quantum Cuts methods for SOD
    - Describe the Probabilistic Saliency Estimation model for SOD.
    - What is a Fully Convolutional Network and how it is used for object segmentation?
    - Describe how the Quantum Cuts method is extended for supervised SOD using Convolutional Kernel Network.

13. Human Action Recognition

- Reading material
  - "Space-time Interest Points", International Conference on Computer Vision, 2003 (download from here: http://www.irisa.fr/vista/Papers/2003_iccv_laptev.pdf)
  - "Action Recognition by Dense Trajectories", Computer Vision and Pattern Recognition, 2011 (download from here: https://hal.inria.fr/inria-00583818/document)
  - "Actions as Space-Time Shapes", International Conference on Computer Vision, 2005 (download from here: http://www.wisdom.weizmann.ac.il/~vision/VideoAnalysis/Demos/SpaceTimeActions/SpaceTimeActions_iccv05.pdf)
  - "Free Viewpoint Action Recognition using Motion History Volumes", Computer Vision and Image Understanding, 2006 (download from here: http://perception.inrialpes.fr/Publications/2006/WRB06a/cviu_motion_history_volumes.pdf)
  - "Multi-view human movement recognition based on fuzzy distances and linear discriminant analysis", Computer Vision and Image Understanding, 2012 (download from here: https://www.sciencedirect.com/science/article/pii/S1077314211002074)
  - "View-Invariant Action Recognition Based on Artificial Neural Networks", IEEE Transactions on Neural Networks and Learning Systems, 2012 (download from here: http://ieeexplore.ieee.org/document/6123211/)
  - "Exploiting stereoscopic disparity for augmenting human activity recognition performance", Multimedia Tools and Applications, 2016 (download from here: https://link.springer.com/article/10.1007/s11042-015-2719-x)
  - "Going Deeper into Action Recognition: A Survey", Image and Vision Computing, 2017 (download from here: https://www.sciencedirect.com/science/article/pii/S0262885617300343), Section 3
- Supplemental reading:
  - "Hidden Markov Models Fundamentals" (download from here: http://cs229.stanford.edu/section/cs229-hmm.pdf)
  - "Human action recognition in stereoscopic videos based on bag of features and disparity pyramids", European Signal Processing Conference, 2014 (download from here: http://www.eurasip.org/Proceedings/Eusipco/Eusipco2014/HTML/papers/1569917287.pdf)
  - "A survey on vision-based human action recognition", Image and Vision Computing, 2010 (download from here: https://www.sciencedirect.com/science/article/pii/S0262885609002704)
  - "A survey of video datasets for human action and activity recognition", Computer Vision and Image Understanding, 2013 (download from here: https://www.sciencedirect.com/science/article/pii/S1077314213000295)
- Slides
  - Visual Human Action Recognition
- Exercises
  - Lab14
- Learning objectives
  - Describe the various sub-problems in human action recognition – HAR (depending on the number of views and the complexity of the scene)
  - Describe the three types of HAR methods, i.e. using global, and local video information and neural networks.
  - Describe various ways to combine human body poses (binary body silhouettes) for representing actions (in one and multiple views)
  - Describe how the bag of features model is used in HAR problem. Consider both the case where global human body information, or local image information is used.
  - Describe Space-Time Interest Points (STIP). How are they used for describing an action?
  - Describe the process followed to calculate the Dense Trajectory-based action description.
  - Describe how a Convolutional Neural Network can be used for HAR when actions are already localized (input to the CNN is a low-resolution video centered to the action). What type of convolutional filters are used?
  - Describe conceptually how we can use neural networks designed for time series in HAR.