UNIVERSITÄT HEIDELBERG
Zukunft. Seit 1386.
IWR Interdisciplinary Center for Scientific Computing

Object Recognition & Image Understanding
Sommersemester 18
Prof. Björn Ommer
ommer@uni-heidelberg.de

Exercise sheet 6: Project Proposal
Mandatory

Due on ***Friday, 8th of June 2018***, 13:00.
Uta Büchler (uta.buechler@iwr.uni-heidelberg.de).

The task of this exercise is to plan your own project and write a small project proposal. To hand in this exercise is mandatory for accomplishing the project and therefore, a requirement for a graded or ungraded certificate. Your description is part of your project and will also be taken into account for your final grade. Therefore we grant you 3 more days for editing this exercise sheet. Please utilize the time to propose a project that you will also be able to accomplish.

General Project Information

The goal of the projects is to give you hands-on experience with more realistic problems compared to the exercises. Therefore, you should propose your own project, where the central part is the implementation of an algorithm for solving a computer vision problem. Besides defining a task you should also name expected results and the performance you want to reach with your algorithm. You are allowed and suggested to work on the project in groups with up to 2 people. But each group member has to provide a substantial contribution to the project and you have to explicitly note who is responsible for which part. For finding teammates you are welcome to use the exercise forum of this sheet. After submitting your project proposal we will check the description and ask you for modifications if necessary. Each student/project group with an approved project plan will be assigned to a supervisor and has to present their work at a poster session at the end of the semester. We will inform you about the proposal until the 13th of June via e-mail. One week before your poster session you have to send your final poster to your supervisor. Please notice that if you are working in a group, each member has to be present at the poster session. For successfully finishing the project you have to send your supervisor your final code and a 5-10 pages long report (per person). Further information regarding the final report will follow. The submitted code should be clear, commented and executable on your supervisor's system. The final grade is based on several evaluation criteria like the integrity of the project proposal, project results (compared to the expected outcome), presentation of the results (poster session), reproducibility of the results, cleanliness of the code and the completion of the report.

Prof. Björn Ommer
ommer@uni-heidelberg.de

Question 1: Project Description(mandatory)
Write a project proposal of 1-2 pages. If you are working in a project group (what we strongly recommend), you should work with your teammates to design only one joint proposal, but each participant has to submit it separately via Moodle. The project description should include the following points.

- **Team:** List the name of all members working on this project (**max. 2 people**). If you are working in a group, please note that each member should have a substantial contribution to the project. Therefore, clearly describe which person is responsible for which part and explain the objectives and expected results of each contribution.
- **Problem Definition:** Describe the task you want to solve and explain why this is interesting. Make sure that the task is not too common and the goal is not too easy (e.g. obtaining 50% accuracy on MNIST classification) and not too hard (e.g. reaching accuries better than state-of-the-art).
- **Dataset:** Please describe which training/test data you want to use for learning based tasks. Try not to spend too much time on this part. We recommend to use or modify existing datasets. See the attached list for some inspiration. If you use an existing dataset please clearly cite the source and mark respective parts in your code.
- **Approach:** Please describe the challenges of your task and classify the problem you are trying to solve using computer vision concepts (at a high level, since we didn't cover all concepts yet). How can you model this task and which approach do you want to follow in your project and why is this appropriate? In the case that you want to use or improve an existing implementation, clearly note this and explain your substantial extension or modification.
- **Evaluation & Expected Results:** Please explain how do you plan to evaluate your algorithm, which performance measure do you want to use and why is this suitable? Also, explicitly state which performance you try to achieve and what the expected results of your approach are. Think about additional properties of your algorithm, how could you design experiments to investigate them and what kind of behavior would you expect?
- **Hardware:** Consider your hardware while designing your project. If you are not sure if it is sufficient, please let us know which hardware you plan to use for the project.

- **Excluded Presentation Date:** Please let us know if you have a serious reason that you can't make it to one of the poster sessions at the 23rd or 24th of July (both from 14.15-15.45). After the assignment to a poster session we can't reassign you to another date due to the large number of students.

Project examples
In the following we list a few problems tackled by the computer vision community (Problems) and show some supervised or unsupervised methods/features which can be utilized to solve them (Methods).

1. Problems
   - object recognition/segmentation/classification
   - image/object matching/registration
   - action recognition (videos)
   - motion analysis, tracking
   - grouping
   - scene understanding
   - generating images
   - face recognition or person identification
   - behavior analysis
   - pose detection
   - ...
2. Methods
   - Learning:
     - support vector machines (SVM),
     - principal component analysis (PCA),
     - clustering,
     - CNNs,
     - LSTMs,
     - GANs,VAEs,
     - reinforcement learning
   - Representations:
     - dictionaries,
     - superpixels,
     - histogram of oriented gradients (HOG),

  - scale-invariant feature transform (SIFT),
  - optical flow,
  - edges,
  - hough.

Following the given examples above possible projects could be

- implementation of an object recognition or classification system,
- extraction/recognition of motion or actions from sport videos,
- visual detection for autonomous driving (e.g.road signs),
- visual learning from synthetic data (e.g. from game environments),
- detailed studies of features and learning algorithms on various datasets,
- segmantic segmentation (e.g. of street scenes),
- image restoration,
- image synthesizing.

Learning Datasets
In the following we list some datasets which you can use for a project on learning. Most of them might be too large for your project but you can use a small subset or only a few categories.

- CV Datasets[1]: A large collection of various computer vision datasets sorted by task.
- Tiny ImageNet[2]: Smaller version of the original ImageNet classification challenge.
- Kaggle[3]: Platform for various machine learning competitions which are uploaded by companies or users.
- Labeled Faces in the Wild [4]: database of face photographs
- MNIST[5]: Large database of handwritten digits.
- CIFAR 100[6]: A labeled subset of the 80 million tiny images dataset.
- Caltech 256[7]: Object category dataset with 256 object categories.

[1] http://www.cvpapers.com/datasets.html
[2] https://tiny-imagenet.herokuapp.com
[3] https://www.kaggle.com/
[4] http://vis-www.cs.umass.edu/lfw/
[5] http://yann.lecun.com/exdb/mnist/
[6] https://www.cs.toronto.edu/~kriz/cifar.html
[7] http://www.vision.caltech.edu/Image_Datasets/Caltech256/

- MIT Places[8]: Scene recognition dataset.
- COCO[9]: A large-scale object detection, segmentation, and captioning dataset.
- UCF101[10]: Action recognition dataset of realistic action videos.
- Cityscapes[11]: Dataset for semantic understanding of urban street scenes.
- KITTI[12]: Street scenes recorded in Karlsruhe.

FYI: Since there won't be the necessity of presenting any solutions on the next exercise session (5.6.) we are going to give you an introduction to the deep learning library *PyTorch* [13] which you will need for solving later exercises on deep learning.

Note: Submit exactly one PDF file via Moodle before the deadline.

[8] http://places.csail.mit.edu/
[9] http://cocodataset.org
[10] http://crcv.ucf.edu/data/UCF101.php
[11] https://www.cityscapes-dataset.com/
[12] http://www.cvlibs.net/datasets/kitti/index.php
[13] https://pytorch.org/