Course Form for PKU Summer School International 2018

<table>
<thead>
<tr>
<th>Course Title</th>
<th>Machine Learning in Computer Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher</td>
<td>Carlo TOMASI</td>
</tr>
<tr>
<td>First day of classes</td>
<td>July 2, 2018</td>
</tr>
<tr>
<td>Last day of classes</td>
<td>July 15, 2018</td>
</tr>
<tr>
<td>Course Credit</td>
<td>2 credits</td>
</tr>
</tbody>
</table>

**Course Description**

**Objective**

Machine learning methods have been employed for a long time in computer vision, but their nature and impact on the field has changed dramatically since 2012, with the introduction of deep networks. This course compares samples of pre-2012 and post-2012 machine learning techniques in the context of image recognition, person detection, human body modeling, and optical flow estimation. After an introduction to basic methods of image processing, including the analysis of optical flow, the course surveys decision trees, random forests, and deformable-parts models as a sample of pre-2012 methods, and reviews their use for person detection and human body tracking. Deep convolutional neural networks are then discussed in some detail, including convolutional pose machines and the OpenPose body tracking system. The course closes with a review of supervised and unsupervised deep learning methods for optical flow estimation.

**Pre-requisites /Target audience**

Senior undergraduate and graduate students with a good grasp of linear algebra, probability, and multivariate calculus. Some light-weight programming may be required for some of the assignments, in a language chosen by the students.

**Proceeding of the Course**

Ten 3-hour lecture sessions and a final 2-hour exam. Readings are a mix of class notes and papers from the literature. Each session is divided into three one-hour sections with brief breaks in-between.

**Assignments (essay or other forms)**

Daily reading assignments and short homework assignments with exam-style problems.
### Evaluation Details
Attendance 20%; homework 45%; exam 35%

### Academic Integrity (If necessary)

### CLASS SCHEDULE
(Subject to adjustment)

<table>
<thead>
<tr>
<th>Session</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1: <em>Overview and Image Processing</em></td>
<td>July 2</td>
</tr>
<tr>
<td><strong>【Description of the Session】</strong> (purpose, requirements, class and presentations scheduling, etc.) Overall structure and contents of the course. Elements of image processing, including convolution, filtering, image differentiation, and image pyramids.</td>
<td></td>
</tr>
<tr>
<td><strong>【Questions】</strong></td>
<td></td>
</tr>
<tr>
<td><strong>【Readings, Websites or Video Clips】</strong></td>
<td>Class notes</td>
</tr>
<tr>
<td><strong>【Assignments for this session (if any)】</strong></td>
<td>Read materials for sessions 1 and 2 (due in session 2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Session 2: <em>Optical Flow</em></th>
<th>July 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>【Description of the Session】</strong> (purpose, requirements, class and presentations scheduling, etc.)</td>
<td>Two-frame image motion as optical flow: Definitions and variational estimation methods for small displacements. The Shift-Invariant Feature Transform (SIFT). Large Displacement Optical Flow (LDOF).</td>
</tr>
<tr>
<td><strong>【Questions】</strong></td>
<td></td>
</tr>
<tr>
<td><strong>【Readings, Websites or Video Clips】</strong></td>
<td></td>
</tr>
</tbody>
</table>
Session 3:  *A Machine Learning Sampler*  
**Date:** July 4  

**Description of the Session** (purpose, requirements, class and presentations scheduling, etc.)  
Basic concepts of machine learning, including decision trees and random forests as examples of “classical” machine learning methods.

**Questions**

**Readings, Websites or Video Clips**  
- Class notes

**Assignments for this session (if any)**  
Read materials for sessions 3 (due in session 4)

Session 4:  *Person Detection and Body Tracking*  
**Date:** July 5  

**Description of the Session** (purpose, requirements, class and presentations scheduling, etc.)  
Methods for finding people in images, and determining their body poses: Histograms of Oriented Gradients (HOG), pedestrian detection, deformable models of body parts.

**Questions**

**Readings, Websites or Video Clips**  
- Class notes

**Assignments for this session (if any)**  
Read materials for sessions 4 and 5 (due in session 5)
Session 5: **Convolutional Neural Networks**

**Date:** July 6

**Description of the Session** (purpose, requirements, class and presentations scheduling, etc.)

**Questions**

**Readings, Websites or Video Clips**
Class notes

**Assignments for this session (if any)**
Homework 2 (due in session 6)

---

Session 6: **Training Convolutional Networks. AlexNet.**

**Date:** July 9

**Description of the Session** (purpose, requirements, class and presentations scheduling, etc.)
Back-propagation of gradients through a deep network, and the use of stochastic gradient descent for training. A description of AlexNet, the first broadly successful deep network in computer vision.

**Questions**

**Readings, Websites or Video Clips**
- Class notes.

**Assignments for this session (if any)**
Read materials for sessions 6 and 7 (due in session 7)

---

Session 7: **The State of the Art in Detection and Body Tracking**

**Date:** July 10

**Description of the Session** (purpose, requirements, class and presentations scheduling, etc.)
Recent deep networks achieve better-than-human performance in object detection tasks, and very high performance for body tracking. Two such systems are reviewed, based on the concepts of batch normalization and convolutional pose machines, respectively.
### Questions

### Readings, Websites or Video Clips

### Assignments for this session (if any)
Read materials for session 8 (due in session 8)

<table>
<thead>
<tr>
<th>Session 8: Supervised Deep Networks for Optical Flow</th>
<th>Date: July 11</th>
</tr>
</thead>
</table>

#### Description of the Session (purpose, requirements, class and presentations scheduling, etc.)
Optical flow is a regression problem, and contracting-expanding networks have proven up to this task. Foundational concepts are reviewed, and a state-of-the-art convolutional flow estimation method is studied.

### Questions

### Readings, Websites or Video Clips

### Assignments for this session (if any)
Homework 3 (due in session 10)

<table>
<thead>
<tr>
<th>Session 9: Unsupervised Optical Flow Estimation</th>
<th>Date: July 12</th>
</tr>
</thead>
</table>

#### Description of the Session (purpose, requirements, class and presentations scheduling, etc.)
Training an optical flow estimator requires large amounts of data. Synthetic datasets and data augmentation are two techniques used to address this difficulty. Another one is to learn flow in an unsupervised manner. Tradeoffs and techniques are studied in this session, including spatial transformer networks and an unsupervised flow estimation system.
### 【Questions】

#### 【Readings, Websites or Video Clips】
- Class notes

#### 【Assignments for this session (if any)】
Read materials for session 9 (due in session 10)

<table>
<thead>
<tr>
<th>Session 10: <strong>Review and Concluding Remarks</strong></th>
<th>Date: July 13</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>【Description of the Session】</strong> (purpose, requirements, class and presentations scheduling, etc.)</td>
<td>This section reviews what was covered in this course and what was not. The future of machine learning and computer vision is discussed. Questions are answered in preparation for the final exam.</td>
</tr>
<tr>
<td><strong>【Questions】</strong></td>
<td></td>
</tr>
</tbody>
</table>

| **【Readings, Websites or Video Clips】** | |

<p>| <strong>【Assignments for this session (if any)】</strong> | Prepare for the final exam. |
| **Session 11: <strong>Final Exam</strong> | Date: July 15 |
| <strong>【Description of the Session】</strong> (purpose, requirements, class and presentations scheduling, etc.) | The final exam takes two hours. It is closed book, closed notes, and covers the main definitions and concepts through questions and simple problems. |
| <strong>【Questions】</strong> | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>【Readings, Websites or Video Clips】</td>
<td></td>
</tr>
<tr>
<td>【Assignments for this session (if any)】</td>
<td></td>
</tr>
</tbody>
</table>