CSC2548: Machine Learning in Computer Vision

Introduction

Sanja Fidler

January 10, 2018
Instructor Info

- **Instructor:** Sanja Fidler (fidler@cs.toronto.edu)
- **Office:** 386 in Pratt
- **Office hours:** Send email for appointment

This course has no TAs, so please bare with me!
Course Information

- **Class time:** Wed at 12-2pm
- **Location:** SS 1070
- **Class Website:**
- The class will use Piazza for **announcements** and **discussions**:
  [piazza.com/utoronto.ca/winter2018/csc2548/home](piazza.com/utoronto.ca/winter2018/csc2548/home)
- Your grade will **not depend on your participation on Piazza**
Course Prerequisites

Good to know:

- Basics of Machine Learning, Neural Networks

Otherwise you’ll need some reading
Requirements and Grading

- This course is a seminar course. We’ll be reading papers on computer vision, covering various ML techniques. Thus, how much you learn greatly depends on how prepared everyone comes to class.

- Each student expected to write short reviews of two papers per week, present a paper/topic, and do a project

- **Grading**
  - Participation (attendance, participation in discussions, reviews): 15%
  - Presentation (presentation of papers in class): 25%
  - Project (proposal, final report): 60%
Project

Logistics:

- Need to hand in a report and do a presentation
- Can work individually or in pairs

Types of projects:

- Great project (A+): nice new research. Does not need to be fully tested by time of presentation
- Good result on a popular benchmark
- Can also implement an existing paper (max grade A, depending how challenging the method is)
- Simply running existing code is not sufficient
## Term Work Dates

<table>
<thead>
<tr>
<th>Term Work</th>
<th>Due Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reviews</td>
<td>one day before class (Tue)</td>
</tr>
<tr>
<td>Project Proposal</td>
<td>Feb 20</td>
</tr>
<tr>
<td>Project Report</td>
<td>end of April</td>
</tr>
<tr>
<td>Project Presentation</td>
<td>end of April</td>
</tr>
</tbody>
</table>

- All dates are for 2018
Lateness

**Deadline**  Reviews / project should be submitted **by 11.59pm on the date they are due**. Anything from 1 minute late to 24 hours will count as **one late day**.

**Lateness**  Each student will be given a total of **3 free late days**. After you have used the 3 day budget, each late day will have a 10% penalty.

**Discount**  You have a budget of 1 missing review without penalty. You do not need to do reviews for the week you present.
Focus on Deep Learning

- Convolutional Neural Networks
- Recurrent Neural Networks
- Graph Neural Networks
- Reinforcement Learning
- Variational autoencoders, GANs
- Graphical models
Computer Vision

Topics:
- Object detection
- Semantic and instance segmentation
- Stereo, flow
- Action recognition
- Tracking
- 3D scene understanding
- Captioning, VQA, retrieval
- Image/video generation, style transfer

How:
- Overview of topic
- We’ll try to cover some old techniques (even if no learning)
- And some of the latest ones
Benchmarks, Resources

Cityscapes: Semantic and instance segmentation
https://www.cityscapes-dataset.com
Benchmarks, Resources

**PASCAL**: Semantic segmentation, detection; 10K images, 20 object classes
http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html

**ADE20k**: Semantic segmentation; 20K images, 150 classes, open voc
http://sceneparsing.csail.mit.edu/
MS-COCO: Detection, segmentation, keypoints, captioning, VQA; 200K images, 80 object classes [http://cocodataset.org/]
Visual Genome: VQA, relationship prediction, attributes, detection...
http://visualgenome.org/
**Benchmarks, Resources**

**KITTl**: Detection (2D, 3D), stereo, flow, tracking, road, odometry

http://www.cvlibs.net/datasets/kitti/index.php
Benchmarks, Resources

**Sintel**: Flow, [http://sintel.is.tue.mpg.de/](http://sintel.is.tue.mpg.de/)

![MPI Sintel Flow Dataset](image_url)
SceneNN: RGB-D segmentation

http://people.sutd.edu.sg/~saikit/projects/sceneNN/
**Matterport3D**: RGB-D segmentation, depth estimation

[https://matterport.com/blog/2017/09/20/announcing-matterport3d-research-dataset/](https://matterport.com/blog/2017/09/20/announcing-matterport3d-research-dataset/)
**House3D**: Room navigation, grounded VQA

https://github.com/facebookresearch/House3D
**Something Something**: Action classification

[https://www.twentybn.com/datasets/something-something](https://www.twentybn.com/datasets/something-something)

<table>
<thead>
<tr>
<th>20BN-SOMETHING-SOMETHING-DATASET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of videos</td>
</tr>
<tr>
<td>Training Set</td>
</tr>
<tr>
<td>Validation Set</td>
</tr>
<tr>
<td>Test Set (w/o labels)</td>
</tr>
<tr>
<td>Labels</td>
</tr>
</tbody>
</table>
**Charades**: Activity parsing; 10k videos

http://allenai.org/plato/charades/
### MovieQA: Video-based QA

http://movieqa.cs.toronto.edu/

<table>
<thead>
<tr>
<th>Movie</th>
<th>The Adjustment Bureau</th>
<th>Snatch.</th>
<th>Revolutionary Road</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question</td>
<td>Why does David abandon Elise at the hospital after she sprains her ankle?</td>
<td>Why does a robber tell Franky to buy a gun from Boris?</td>
<td>Why does April die?</td>
</tr>
<tr>
<td>Story</td>
<td><img src="image_url" alt="Image" /></td>
<td>- When you get to London...</td>
<td>April dies in the hospital due to complications following the abortion.</td>
</tr>
<tr>
<td>Correct answer</td>
<td>To protect both Elise and himself from Thompson’s threats</td>
<td>Because the robber and Boris want to steal the diamond from Franky</td>
<td>She performs an abortion on her own</td>
</tr>
<tr>
<td>Wrong answer 1</td>
<td>Because he wants to be with someone that can walk</td>
<td>He wants to hook him up</td>
<td>Due to injuries from an accident</td>
</tr>
<tr>
<td>Wrong answer 2</td>
<td>He wants to run the Bureau and he cannot do it with a limping wife</td>
<td>He plans on robbing and killing him</td>
<td>She kills herself</td>
</tr>
<tr>
<td>Wrong answer 3</td>
<td>He does not abandon her, he stays with her</td>
<td>Because otherwise Boris would kill him</td>
<td>Due to a drug overdose</td>
</tr>
<tr>
<td>Wrong answer 4</td>
<td>He wants to save her from World War I and the Great Depression</td>
<td>The robber plans to steal a painting from Franky</td>
<td>She is shot</td>
</tr>
</tbody>
</table>