COS 598B: Advanced Topics in Computer Science -- Visual Recognition

Olga Russakovsky

Course goals

- Deep dive into computer vision
  - Gain a deeper understanding of the selected topics, including key papers and key players
  - Explore the interplay between CV tasks, datasets, methods, analyses, results
  - Discover entry points to learn more about other computer vision topics of interest
- Develop/reinforce research skills
  - Become comfortable reading research literature, incl. doing a literature search
  - Hone both written and speaking scientific communication
  - Practice fair & constructive peer review

Course structure

- Grad seminar, thus assume largely self-motivated
- Reflects the diversity of students here
  - Can/should be adapted to fit your needs
- 2 components: lectures and project

(in reverse order) Component #2: Project

- Work alone or with a partner
- Produce an 8-page paper in CVPR format
- Provide feedback on projects to 1-2 other teams
- Schedule
  - **Wed, April 11th** in class: project title, selection of option 1-3, (optional) partner name
  - **Fri, April 13th**: milestone report, 4+ pages
  - **Fri, April 20th**: milestone feedback due to your assigned team(s)
  - **Mon, April 30th and Wed, May 2nd:** project spotlights in class
  - **Fri, May 6th**: project report due
  - **Dean's date May 15th**: report feedback due to your assigned team(s)

Project option a) Computer vision system

- Project similar to COS 429 requirements last semester
- Build and analyze a computer vision system
- Report should include
  - Introduction positioning the problem and proposed approach
  - Overview of related work
  - Description of the method
  - Quantitative and qualitative results
  - Analysis and conclusion

Project option b) Analysis

- Pick a visual recognition topic and perform an analysis of existing techniques
  - Decide on a particular angle of analysis (across datasets, across methods, etc)
  - Download the related code/data/annotations, code/script to generate the results
- For inspiration, check out
  - [Toralba and Efros ICCV'11] on image classification datasets
  - [Russakovsky et al. ICCV'13] on large-scale object localization
  - [Sigurdsson et al. ICCV'17] on human activity recognition
- Report should include
  - Introduction positioning the investigation, related work, lots of analysis and conclusions

Project option c) Literature review

- Think of this as a book chapter
- Can be about a topic from class but much more in depth
  - Should include 20+ citations, both classical literature and latest techniques
- Target audience: COS 429 student who wouldn't take COS 598B
- See e.g., Crowdsourcing in Computer Vision
  - But yours will be 8 pages, with fewer citations but more in-depth look at each one

Component #1: Lectures

- 3 modules: (1) pixel-level understanding, (2) language+vision, (3) video analysis
  - Each module is 3-4 weeks, thus 6-8 lectures, with 1-3 papers per lecture
  - Generally lectures earlier in a module are predetermined by papers we need to cover, and later are more flexible and can be guided by your interest
- Most lectures given by you
  - Assume little hand-holding on logistics-- but happy to provide all the help you want on content
  - Will get feedback from classmates afterwards
  - Can lecture in pairs if you prefer but then expect do to ~2x more lectures

When giving a lecture

- Come meet with me before your lecture
  - Wed 4:30-5:30pm in CS 408, or by appointment
  - For more junior students: 2+ weeks ahead of time (!) recommended
  - Helpful to have read the paper and drafted a rough plan before the meeting
- Coordinate with me & post on Piazza if there's a paper (or a section of a paper) that would be helpful as background reading
- Post slides after the lecture
- Take responsibility for figuring out a backup plan if you're not going to be there suddenly

When giving feedback on a lecture

- Expect to do this ~3x more times than lectures
- Must read the papers beforehand
- 1+ page feedback emailed to the present(s), cc'ing me, within a day of the presentation
    - Comment on clarity, completeness, slides, …
    - Offer constructive criticism but also suggestions

When not giving the lecture or feedback

- Come to class (duh)
- Ask lots of questions
- No email, faceboook, twitter, snapchat, ...
- Good idea to read the papers beforehand
  - Definitely read the background papers, if posted
- If you have to miss class: read the papers and look at slides to catch up

Let’s look at the schedule

- Schedule is here, and will be updated throughout the semester
  - Presenters will post links to slides, background readings, papers, etc.
  - Please *comment* on the doc with updates and I’ll incorporate the changes into the main text

Week 6 (March 12-16): Option 1

- Class as in, add another week of e.g., video analysis
- **Must commit to coming** despite midterms & ECCV deadline

Week 6 (March 12-16): Option 2

- Class canceled in favor of midterms & ECCV deadline
- Instead: attend the 3 amazing computer vision CS department seminars later in the spring
- Also: PhD students, help extra with Visit Day March 15-16

Sign up for lectures

- Fill out this Google form to submit your preferences
- Volunteers for week 2: FCN and FCN weakly supervised
- Any questions on topics, structure, …?