
COGS 260: Final Project

1 Overview

The Course Project is an opportunity for you to apply what you have learned in class to a problem of your interest. You are encouraged to select a topic and work on your project. Potential projects usually fall into these two tracks:

- **Applications:** Pick a real world problem and use ConvNets or RNNs to solve it.
- **Models:** You can build a new model (algorithm) with ConvNets, RNNs, or a new variant of existing models, and apply it to tackle vision tasks. This track might be more challenging, and sometimes leads to a piece of publishable work.

Some of the interesting topics that you can work on are listed below:

- Image -to-sentence generation [Karpathy and Fei-Fei] [Donahue et al.] [Vinyals et al.]
- Music composition [Eck, D. and Schmidhuber, J.] [Boulanger-Lewandowski et. al.]
- Object recognition [Krizhevsky et al.] [Russakovsky et al.] [Szegedy et al.][Simonyan et al.][He et al.]
- Object detection [Girshick et al.] [Sermanet et al.] [Erhan et al.]
- Image segmentation [Long et al.]
- Video classification [Karpathy et al.] [Simonyan and Zisserman]
- Face recognition [Taigman et al.]
- Visualization and optimization [Szegedy et al.] [Nguyen et al.] [Zeiler and Fergus] [Good-fellow et al.] [Schaul et al.]
- Deep Dream [Google Deep Dream]
- Neural Style [JC Johnson]

Here are a list of popular computer vision dataset:

- Meta Pointer: A large collection organized by CV Datasets.
- Yet another Meta pointer
- ImageNet: a large-scale image dataset for visual recognition organized by WordNet hierarchy
- SUN Database: a benchmark for scene recognition and object detection with annotated scene categories and segmented objects
- Places Database: a scene-centric database with 205 scene categories and 2.5 millions of labelled images
- NYU Depth Dataset v2: a RGB-D dataset of segmented indoor scenes
- Microsoft COCO: a new benchmark for image recognition, segmentation and captioning
- Flickr100M: 100 million creative commons Flickr images
- Labeled Faces in the Wild: a dataset of 13,000 labeled face photographs

- Human Pose Dataset: a benchmark for articulated human pose estimation
- YouTube Faces DB: a face video dataset for unconstrained face recognition in videos
- UCF101: an action recognition data set of realistic action videos with 101 action categories
- HMDB-51: a large human motion dataset of 51 action classes

2 Team

The project can be done in a group of at least 2 and a maximum of 3 students if you are using the GPU resources provided by us. If you have your own resources then you are allowed to do even individual project.

3 Important Dates

- **Report Due Date:** June 9 (Thursday), 11:59 PM, 2016. To be submitted on TritonEd.
- **Late policy:** A grace period of 3 hours will be given, which means no penalty assigned if you turn it in no later than 12/11/2015, 2:59 AM. A penalty of 5% of the total points will be deducted for every extra day past due.
- **Presentation:** June 10 (Friday), Time: TBD

4 Grading Policy

The merit and grading of your project can be judged from aspects described below that are common when reviewing a paper:

- Interestingness of the problem you are studying. (10 points)
- How challenging and large is the dataset you are studying? (10 points)
- Any aspects that are new in terms of algorithm development, uniqueness of the data, or new applications? (20 points)
- Is your experimental design comprehensive? Have you done thoroughly experiments in tuning hyperparameters? (30 points)
- Is your report written in a professional way with sections including abstract, introduction, data and problem description, method description, experiments, conclusion, and references? (30 points)
- Final presentation. (15 points)
- Can you demo your work through a website or an app during the final presentation? (5 points)
- Bonus points (upto 20 points) will be assigned to projects that have adopted more than one evaluation metrics, tested on more algorithms, tried new methods, and worked on novel applications.

5 Report

Write a report with more than 1500 words in any leading machine learning journal format like NIPS and ICML. Your project report should contain the following sections:

- Title, authors, emails of authors.
- Abstract of the report. The last couple of lines of the abstract should state the main result, e.g., "We achieved 76% correct on the so-and-so dataset."
- Introduction and motivation: Why is this interesting? There should be reference to prior work, especially work you are replicating, if that's what you are doing.

- Methods: Data (including train and test split), Model (design of the network), training procedures (details like, how you initialized the weights, what your objective function was, learning rates, etc.), testing procedures.
- Results: Nicely labeled graphs of your results with appropriate text describing them.
- Discussion and Conclusions
- Acknowledgements (optional): If you received a lot of guidance from your TA or other graduate students, state that here.
- References: Nicely formatted references in APA format or NIPS format.

6 Sample Projects

Some sample projects can be found on [Stanford CSE 231n Project Page](#).