Final Project: Classification of Visual Information

Teams formed and approved by Tuesday, March 5 (end of class)
Project Proposal due by Thursday March 21 (start of class)
Project Proposal Approved by Thursday March 28 (end of class)
Final Project Report due Monday April 29 (midnight)
Project Presentation due May 2 (midnight)
Team Presentation Required: Friday, May 3, 1:30-4:00pm, Lewis Ross Hall (ROS) A310

CMPE 685 Teaching Assistant: Nilesh Pandey np9207@rit.edu

Objectives
The objectives of the final project are:

- Use deep neural networks for a computer vision task involving detection, recognition, segmentation, or other computer vision topics. For example: object detection (faces, people, cats, dogs, cars, road signs, etc.), face recognition, expression recognition, action recognition, scene classification, semantic segmentation, object tracking, etc.
- Train and test deep network architectures on multiple datasets.
- Review current literature (journal or conference publications) relevant to your project.
- Collaborate with a teammate to complete the project and write the project report.
- Present your project in class for the benefit of your classmates.

Project Proposal (10%):
A project proposal is required. A pdf version of the project proposal (pdf and zipped overleaf project) should be uploaded on mycourses by the due date/time. A printed proposal should be submitted in class. The project proposal should include: Project Title, Project Team Members, Project Goals, Related work, Methodology including network models, development environment, datasets (with URL) and anticipated results, a Time Table with Milestones and Work Assigned to each team member during each phase of the project.

Project Report (70%):
A pdf version of the report (pdf and zipped overleaf project), images (zipped), and all code (zipped) should be uploaded on mycourses by the due date. The submitted project report in paper form (without the code) is due at the start time of the final exam. Your report should include the following sections: Introduction (include problem statement), Related Work (discuss at least two relevant papers, one per team member, in current literature), Methods (network architectures, parameter selection, classifier, etc.), Experiments (datasets, experimental protocol), Results (with tables and examples of your system's performance), Discussion (comparisons, advantages, limitations, future work), Teamwork, References (include URL's for datasets or code). **In the teamwork section, include a table that clearly states the contribution of each team member for each part of the report and for each architecture, dataset and result.** If the team members collaborated on a section or to get a result, the approximate level of effort should be included (e.g. 50% A and 50% B, or 80% A and 20% B). Use color figures in the pdf version of the report and grayscale images in the figures of the printed version. Make sure you maintain the image aspect ratio when displaying images in the figures. The report will be graded based on reviewing related work, methods used including architectures and datasets, experiments, performed, results, interpretation of results, clarity of presentation, discussion and English.

Project Presentation (10%):
A project presentation by both members of the team is required during the final exam time. The pdf of the presentation should be uploaded on mycourses by the due date/time. Do not upload a presentation in any other format except for pdf. The presentation should take 5 mins and have 8-12 slides: title (1 slide), Problem statement (1 slide), Related Work (1-2 slides), Methods (2-3 slides), Experimental Methods (1-2 slides), Results (1-2 slides), Critical Evaluation (1-2 slides), Conclusions (1 slide), References should be placed at the bottom of the slide where they are cited. All students are expected to attend the presentations and participate in the evaluation of each presentation. The evaluation criteria for the presentation are: Quality of slides, Clarity, Presentation skills, Within Time Limits (5 min).

Teamwork (10%):
Each student will be evaluated on their contribution to their team and their project, including methods and results, writing the proposal and report, and presentation.

Project Requirements
Work with a classification system that has the following components:

a) You are required to use a **deep learning architecture**, e.g. Convolutional Neural Network (CNN). Each team member should use a different architecture for which they will be responsible. You may use one of the existing architectures (recommended) or design your own network (needs to be trained from scratch). If your target dataset is relatively small, you can perform **transfer learning** on a pre-trained model instead of retraining the full network. Some popular architectures are listed below, but you are allowed to utilize other models.
b) **At least two datasets**. Some commonly used datasets are listed below, but you are allowed to use other datasets (include the dataset descriptions in your project proposal)
c) *Optional:* **Preprocessing** of the data, e.g. **data augmentation**, may be used to increase diversity in the training data, by introducing variations in orientation, size, illumination, etc.
d) **Classifier**: select the type of classifier that you will use with your network, e.g. Fully Connected network (select number of layers) or SVM.
e) **Training and testing methodology**: Decide on the training and testing process, e.g. select k for k-fold cross validation. Make sure the classifier is tested on a test dataset that is completely different from the training set. The objective is to design a classifier that is accurate and has the ability to generalize, i.e. deal with data outside the training set. Discuss ways to prevent overfitting on the training dataset.
f) **Hyperparameters**: Explore various settings of the network parameters, e.g. learning rate, batch size, etc. Additionally, you may experiment, with different loss functions, nonlinearities, etc.
g) **Comparisons**: Compare your results for: at least two different architectures and two different datasets. If applicable, compare your results with other methods found in literature.
h) **Teamwork**: the work is typically done by a team of two students that equally divide the tasks involved in the design, training and testing of the computer vision system. A small number of advanced students may request individual projects. If you want to do an individual project, you should request it as soon as possible.

Image Datasets

It is recommended to use standard datasets used in Computer Vision and Image Processing. Many links to datasets are found at:
http://homepages.inf.ed.ac.uk/rbf/CVonline/Imagedbase.htm
http://riemenschneider.hayko.at/vision/dataset/
http://www.cvpapers.com/datasets.html

Popular Network Architectures

- AlexNet https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks
- VGG http://www.robots.ox.ac.uk/~vgg/research/very_deep/
- GoogleNet https://ai.google/research/pubs/pub43022
- ResNet https://arxiv.org/abs/1512.03385
- DenseNet https://arxiv.org/abs/1608.06993
- Faster R-CNN https://arxiv.org/abs/1506.01497
    - Fast R-CNN https://arxiv.org/abs/1504.08083
    - R-CNN https://arxiv.org/abs/1311.2524
- Mask R-CNN https://arxiv.org/abs/1703.06870
- Yolo https://arxiv.org/abs/1612.08242
- SqueezeNet https://arxiv.org/abs/1602.07360
- and more…