Computer Vision **Name:** _____
Prof. Bolton **Net ID:** _____
Final Project

This assignment contains 4 pages (including this cover page) and 1 questions. Total of points is 100.

Conditions: Work in groups of two or less.

Include your name and Net ID. Follow submission instructions as indicated herein.

---

1. (100 points) For this assignment you will complete a large scale project with applications to computer vision. There are multiple pre-determined options noted; choose one of these options ... or, if you wish to attempt a final project that is not on the list, that may be OK with instructor approval. Note: All projects will require a brief proposal which details the project scope. Final projects will be presented in class on the final 3 days (approx 15 minutes each).

1 Proposal

All proposals should be sent to me via email and should include the following information:

1. Name of individual or names of group members
2. Topic Choice and/or title
3. Scope detail. Note: there are many pre-existing methods available in Python and Matlab to perform various Computer Vision tasks. You may use some of these existing methods. However, you cannot use existing methods that perform the "crux" of your project. Thus, please clearly identify what code you will be creating using your own design. (I will attempt to clearly identify my expections with respect to what pre-existing code is usable and not usable for each project.)
4. Data set for experimentation
5. Preferred presentation date(s) in order of preference (FIFO)

2 Topics

Below is a list of pre-determined topics. You may deviate from this list only with prior approval via a proposal submission and approval. In addition to the task noted within each topic there is an "group task", which is a task to be performed if you are in a group.

2.1 Predetermined topics

For each project, you will implement the corresponding algorithm(s) and apply to a data set for validation and/or experimental comparison.

1. Graph-based Methods. For this option, the crux of the approach is mapping the computer vision problem to the graph domain. Thus the construction of the graphs and weights must be designed and implemented by you. Once the graph is implemented you can use standard max-flow / min-cut algorithms.
   - Disparity Map Estimation: Implement Roy and Cox [1], "A Maximum-Flow Formulation of the N-Stereo Correspondence Problem". *Group: Implement Boykov et al [6] and compare results and efficiency to that of Cox. Note: you must implement both the swap and extension operations.*
   - Image Segmentation. Boykov et al [5] *Group: Implement another image segmentation technique and compare restults to Boykov et al.*
2. Implement an Active Contour (Snake) or Deformable Model, Kass et al. [2]. *Group: Use active contours to aid in tracking an object in a sequence of images (video).*
3. Convolutional Neural Networks. Learn Caffe (http://caffe.berkeleyvision.org/) and experiment on a moderate-sized data set (e.g. http://image-net.org). See recent ACM paper [3]. Perform cross validation experiments where the goal is object classification or something similar. *Group: Compare the results of different CNN configurations OR perform perturbation analysis.*
4. Multiple Instance Learning: applications to computer vision. Maron et al. Implement MIL approach similar to that of Maron et al. for computer vision applications. Try your approach on Correl: Fox, Tiger, Elephant; or similar. [4] *Group: Compare results to some non-MIL object recognition approach.*

2.2 Innovation

Incorporating innovative methods or designs to one of the existing methods may result in "bonus" points to be determined by the instructor. If you attempt to improve or add innovation to existing techniques, I encourage you to attempt this after completing the base portion of the project. Feel free to share any thoughts with me and I can help to guide you through this process.

3 Data Sets and Results

For the sake of comparison, I encourage you to use the data sets used in the original papers. But in general, feel free to use any reasonable (given your topic selection), publicly-available data set.

Qualitative and/or quantitative results should be presented. Example of qualitative results would be detection maps (detection image), segmentation image, ... Examples of quantitative results would be ROC curve, confusion matrix, ...

4 Presentation Format

Presentations should be around 12 minutes + 3 minutes for questions. Presentations should focus on concepts, implementations and results. Given the time constraints, I encourage you to create many diagrams, illustrations, tables, charts and the like so that large amounts of information can be conveyed as efficiently as possible. Please follow the standard format below:

1. Problem Overview (2 minutes)
2. Solution Strategy and Implementation (7 minutes)
3. Experimental Design and Data Sets (2 min)
4. Results, Comparisons and/or Conclusions (1 min)

5 Rubric

List of Requirements	**Percentage**
Comprehensive coverage of topic / scope as stated in proposal	0.20
Correctness and efficiency of implementation	0.40
Experimental design / validation well-conceived	0.10
Presentation well designed and conveyed	0.20
Relative (to peers) performance (by quartile ranking)	0.10
Innovation	α
TOTAL	**1.0** + α

References

[1] Roy, Sébastien and Cox, Ingemar J. A maximum-flow formulation of the n-camera stereo correspondence problem, *Computer Vision, IEEE Sixth International Conference on*. pp 492–499, 1998.

[2] Kass, Michael and Witkin, Andrew and Terzopoulos, Demetri. Snakes: Active contour models, *International journal of computer vision*, vol. 1, no. 4, pp. 321–331, Springer 1988.

[3] Jia, Yangqing and Shelhamer, Evan and Donahue, Jeff and Karayev, Sergey and Long, Jonathan and Girshick, Ross and Guadarrama, Sergio and Darrell, Trevor. Caffe: Convolutional architecture for fast feature embedding, *Proceedings of the 22nd ACM international conference on Multimedia*, pp 675–678, 2014.

[4] Maron, Oded and Ratan, Aparna Lakshmi, Multiple-Instance Learning for Natural Scene Classification, *ICML*, vol. 98, pp 341–349, 1998.

[5] Boykov, Yuri Y and Jolly, M-P, Interactive graph cuts for optimal boundary & region segmentation of objects in ND images, *Eighth IEEE International Conference on Computer Vision (IEEE ICCV)*, vol. 1, pp 105–112, 2001.

[6] Boykov, Yuri and Veksler, Olga and Zabih, Ramin, Fast approximate energy minimization via graph cuts, *IEEE Transactions on pattern analysis and machine intelligence*, vol 23, no. 11, pp 1222-1239, 2001.