CSE/ECE-344/544 - Computer Vision
Course Project Guidelines

Winter 2024

Project Group Size: 3 students

Timelines:
Project description: 11:59PM, Apr. 23, 2024 - 4 slides (see details below)
Final Presentation and Report submission: 3PM, May. 14th, 2024
Final presentation & viva: May 14th and 15th (Timings to be shared).

Grading Break-up:

1. Project Description: 5 points
2. Final Presentation: 12.5 points slides + presentation + viva
3. Final report: 12.5 points (report including paper review and analysis)

Project Areas

1. Visual Recognition
   (a) Object recognition: Benchmarking on *unusual* datasets - Indian driving dataset (IDD), Wildlife dataset, Medical imaging datasets. Error analysis using TIDE.
   (b) Visual Re-Identification of Wildlife using MegaDescriptor.
   (c) Instance / Panoptic Segmentation: Cityscapes dataset, IDD, KITTI, etc.
2. Geometric Vision
   (a) SLAM: Visual SLAM (monocular & stereo) [ORB-SLAM3], LIDAR-SLAM [FLOAM, CT-ICP], Semantic SLAM [Kimera], and their error analysis. You may use autonomous driving datasets that have visual as well as LIDAR data, e.g., KITTI, NuScenes, etc.
   (b) Multi-view Homography Estimation [ProgressiveX].
3. Vision and Language

(a) Visual Question and Answering
(b) More details to come.

Other areas are fine too, so long as you have already started working on them. If you haven't, please select one of the areas listed above.

Project Execution Guidelines

1. Choose one topic among the three areas that are listed above.
2. Pick two relatively recent papers (last 4-5 years) in this area with *both* code and a pre-trained model available. Some papers are listed, and you may choose others. However, please make sure that sufficient compute is available for any inference (not training) that needs to be performed.
3. Reproduce the paper results by performing inference on the *pre-trained model* on one of the datasets used in the papers.
4. Identify another related dataset that is *not used* for training in the paper. Design an analysis experiment using the evaluation metrics to analyze the different types of errors between the two approaches and report your findings.

Submission and Evaluation Details:

1. **Project Description**: Each group has to prepare and submit 4 slides to present the following:
   (a) problem statement
   (b) dataset (one from the paper and the other from elsewhere)
   (c) approaches investigated - find two papers that you plan to read about and investigate.
   (d) evaluation metrics and inference time hardware requirements.
2. **Final presentation**: 7 minute presentation followed by Q&A. More details will be shared soon.
3. **Final report**: 4 page report (in the 2-column CVPR format) describing
   (a) the problem statement
   (b) description of the approaches used; short technical summary of the papers you have used.
   (c) the results of the analysis experiments performed on the two datasets.
   (d) You **can** (and are encouraged to) use LLMs like chatGPT extensively for this part of the report.
   (e) The last section of your report should contain a one-column (i.e., half page) commentary written by each team member independently without using chatGPT or any LLM tool. The commentary must summarize the analysis experiments and your conclusion about the performance difference (if any) between the two papers.