AI3603: Computer Vision, Spring 2025

Indian Institute of Technology Hyderabad
Homework 5, Optical Flow, Stereo Vision, Object Detection
35 points. Assigned 21.04.2025, Due **11:59 pm on 26.04.2025**

If you have integrity, nothing else matters. If you don't have integrity, nothing else matters.. – Alan Simpson

Instructions:

- It is **strongly recommended** that you work on your homework on an *individual* basis. If you have any questions or concerns, feel free to talk to the instructor or the TAs.

- For Q1, use `https://www.bogotobogo.com/python/OpenCV_Python/images/mean_shift_tracking/slow_traffic_small.mp4` and `https://github.com/opencv/opencv/blob/master/samples/data/vtest.avi`.

- For Q2, use the stereo pair uploaded with this HW.

- For Q3, use the `dog-bike-car` image uploaded with this HW.

- You are free to use Copilot. Please turn in your prompts.

- Please turn in Python Notebooks with the following notation for the file name: `your-roll-number-hw5.ipynb`.

1 Optical Flow (15)

1. Compute the dense optical flow of the aforementioned videos using the Horn-Schunck method discussed in class. For brevity, display the flow for the first three frames. (5)

2. Repeat the same using the FlowFormer method. Again, display the flow for the first three frames. (5)

3. Compare the two methods in terms of computational complexity and accuracy. Which one is better? Why? (5)

2 Stereo Vision (5)

Implement a disparity estimator using mean absolute error (MAE) as the distance measure. Use the left and right images upload with this HW. Assume that the images are rectified. Display the disparity map. (5)

3 Object Detection (15)

1. Apply the YOLOv1 model discussed in class to the `dog-bike-car` image and display the results. Use pre-trained weights from a standard implementation. (5)

   (a) Now rotate the image by 90° and repeat. Note your observations.

   (b) Downsample the image by a factor of 2 and repeat. Note your observations.

   (c) Downsample the image by a factor of 4 and repeat. Note your observations.

2. Repeat the above experiment using the RT-DETR mode. (5)

3. Compare the two models in terms of accuracy, speed and robustness. Which one is better? Why? (5)