People detection Laboratory

Session1 (session1.m script)

1-Run **test_ReadGTBlobs** script

- Describe **test_ReadGTBlobs** script (2-3 lines)

  Ans.) This script reads and analyzes 71 images in the TUD-sequences frame. After reading the images, the script calls the function ReadGTBlobs and passes the ground truth and the images sequence to this function and the function returns Blobs for each image. Finally, in the shown output images there are bounding boxes where people are detected in the frame sequence.

- Describe **ReadGTBlobs** and **PaintBlobs** (2-3 lines)

  Ans.) **ReadGTBlobs:-** The function ReadGTBlobs takes the ground truth and the images sequence as the input. Initializes a processing mask, and the determines the X1,X2, Y1,Y2 values. Using these X1,X2,Y1,Y2 values the W,H values and the score is determined for each blob in the Image. The function returns all the determined Blobs for each image in the frame sequence.

  PaintBlobs:- For each image, the function PaintBlobs takes the blobs, the image and color as inputs. For the given frame, for each of the set of values for the blob, it draws the bounding box,where people are detected.

- Describe blobs output format.

  Ans) The Blobs are cells array of size 1x71. Each cell contains a struct(blob), the size of the struct depends on the number of blobs(X,Y,W,H and score) (number of people) in the given image in the frame.

2-**test_DTDP_detector** script

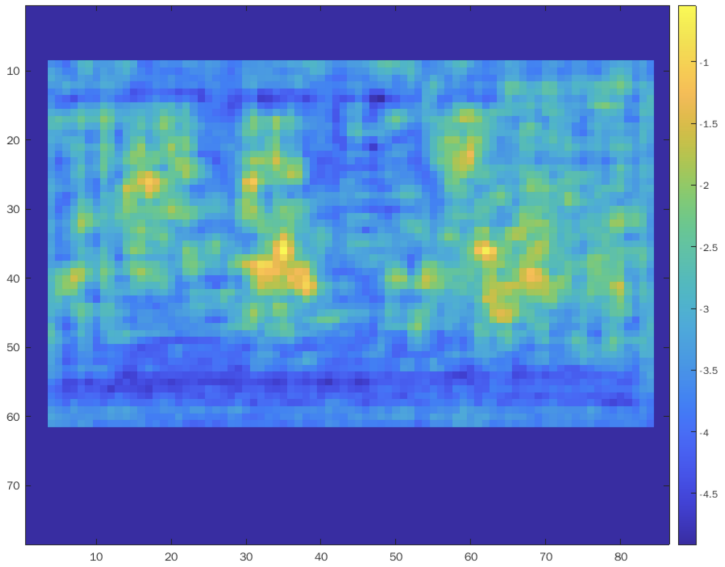
- Directory DTDP detector/....windows32bits/
  - Configure your compiler (run “mex –setup” Matlab command, and follow Matlab intructions)
  - Run **compile** script
  - Run **demo** script
  - Describe **demo** script (2-3 lines)
    Ans.) The demo script loads irina_person.mat file and calls the test function. An image and the model is passed to the test function. The model is then visualized using the visualizemodel function

and corresponding images of different body parts learned by model are displayed. In the original given image, different body parts of the people are detected. Corresponding to the different body parts of people (blue color bounding boxes), the red bounding box are made for detected people.

- Paper "Object-Detection-with.....pdf"
  - Compare results and discuss differences with Figure 1, Figure 1-a, b and c.
    Ans.) The Figure-1 from Demo.m script uses HOG image of person and divides it into 8 parts, to detect the different body parts for the person in the given image. The Figure-1 a,b,c uses different detection model. a, uses root with head for the detection; b, uses root and 5 body parts for detection and c, uses root and 8 different body parts for the detection.
  - Person model:
    - Identify the number of body parts.
      Ans.) As per the demo script,a total of 9 body parts were identified(8 body parts+ root).

      Identify the number of poses supported (mixture models).
      Ans.) As mentioned in the object detection pdf, Twenty(20) different configurations. as per the demo script one pose.

  - Identify in the code where the Image and Feature pyramid in Figure 3 is generated and stored.
    Ans.) The Image and feature pyramid is generated and stored using the function featpyramid. The imgdetect script calls this function.
    - Identify the number of images in the pyramid or scales.
      Ans.) Model.Interval determines the number of features. a total of 45 features were detected.(Scales: 45x1 doubles)
  - Identify where the final combined score is (see figure 4).
    Ans.) – Final combined score is determined by the symbolscore function in gdetect.
    - Visualize (with imshow) the final combined score at scale (level) 11.
      Ans.) used imagesc inplace of imshow to add scaling



- Identify and explain the Non-Maximum Suppression (section 7.2)

  Ans.) In the script, the nms function is called in the demo script at lines 36 and 46. A non-maximum suppression process in used in order to eliminate possible repeated detection. The window with the highest score is considered and all the windows with lower score are dropped off. Non maximum suppression is used for reduction of overlapping detection in the image, where the ratio of the intersection of two overlapping detections to their union is large than the threshold, the detection with the smaller score gets eliminated.

3-Run **test_DTDP_detector** script

- Describe **test_DTDP_detector** script (2-3 lines)

  Ans.) The script loads tud campus image sequence and inriaperson_final.mat as model.Calls DTDP detector and passes images_names,threshold,model id and filename as parameters.
- Describe **DTDP_detector** script (2-3 lines)
  Ans.) The DTDP detector script calls the test function for all the images in the image cell array and passes model, threshold, image sequence. Uses imgdetect to determine the bounding boxes. uses reduceboxes to eliminate the unused filters and calls save_blobs to append the final blobs to a file.
- Describe blobs output format.

  Ans.) The final blob has X1,Y1,X2,Y2 and score values and those values are saved in a *.idl file as X,Y,W,H

4-**test_ReadBlobs** script

- Run **test_ReadBlobs** script
- Describe **test_ReadBlobs** script (2-3 lines)

Ans.) The script uses Inria model, tud-image-sequence and the saved *.idl file for the blob detection in the tud-image-sequence. It calls readblobs function to determine the blobs (X1,Y1,X2,Y2) and their score in the given image frame. Paintblob is used to display the image and bounding boxes where people are detected.

- Select a debug_threshold value in order to avoid as many false-positive detections as possible.
  Ans.) -1.6 value can be used.

- Describe blobs output format.

  Ans.) The blobs is a cell array which contains struct. The size of the struct depends on the number of blobs(X,Y,W,H and score) (number of people) in the given image in the frame.

5-Repeat **demo** script and **steps 3 and 4** with VOC2009 person model instead of Inria person model

- Compare the visual results from both detectors.

  Ans.) In the Demo script, while using the Inria model, it divides the person into 8 parts and for the given Image, detects correctly the 2 persons, where as VOC2009 uses 3 types of features 1. upper part of the body and divide the upper part into 8 parts. 2. Person in a different orientation(sitting maybe) and divides into 8 parts 3. standing person and divides into 8 regions. The results for VOC2009,shows false detection in the Image(lesser accurate results than Inria).