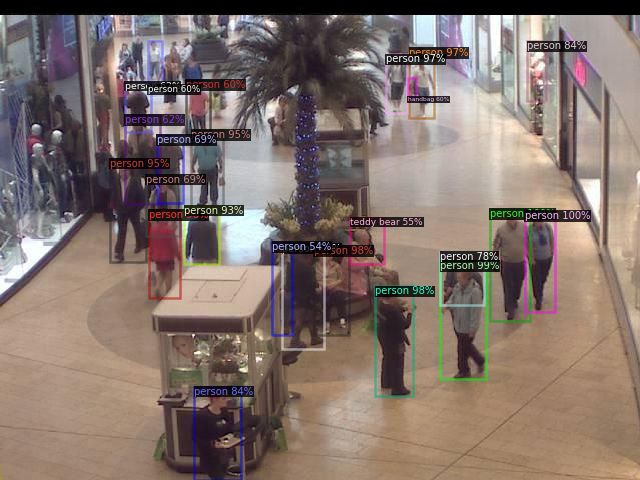
**Description of Person Detector (part 1):**

Todo: where was the method obtained

For the initial part of the project, Detectron2, an object detection platform written on PyTorch has been utilized [1]. Detectron2 offers several models; in this project, the object detection model is used. Here, instances are classified and localized with bounding boxes [2]. Object detection is performed on images in the ‘mall’ dataset to detect people and obtain bounding boxes around them.

The model is pretrained on the COCO dataset. In this case, the faster\_rcnn\_R\_50\_FPN\_3x model is used. Faster R-CNN is composed of two modules: a fully convolutional network and a Fast R-CNN detector. It is a “single, unified network” for object detection [3]. The features of the images are extracted and the regions where person objects exist are found. Regions are then bounded by boxes. The results can be seen in the image below (*Figure 1)*.



*Figure 1*: Using Faster-RCNN to detect person objects

After that, person objects have been cropped out and saved in a separate file. These images will serve as positive examples. Negative examples have been produced by randomly cropping sections of different dimensions. These are illustrated below in figure 2.  


*(a) (b)*

*Figure 2:* (a) Positive example. (b) Negative example.

**Description of Person Detector (part 2):**

Todo: describe dataset number of training examples, explain how LBP and HOG are combined

In this part a person detection algorithm trained with a Support Vector Machine (SVM) classifier is built from scratch. SVM has the highest precision among other classifiers for human detection [5]. SVM is a binary classification algorithm that makes use of an optimal hyperplane as a decision boundary [4]. Histogram of Oriented Gradients (HOG) values are used as the feature sets and SVM is used to train the classifier.

Feature extraction is one of the fundamental goals. The cropped images from part one will be used to train the model. The images are divided into positive examples: images containing a person, and negative examples: images not including a person. HOG feature vectors for the images are found. The positive and negative features are labelled 1 and 0 respectively.

In HOG, the features comprise the pattern in the directions of gradients. Corner and edge detection is fundamental to image detection. The appearance and shape of a local object can be identified reasonably well using the direction of edges [6]. A higher magnitude of edge correlates to a corner or edge. HOG descriptor is calculated by finding the horizontal and vertical gradients. The image is divided into cells, for each cell a histogram of the directions of the gradients is made. Accumulation of the histograms forms the HOG representation [6].

A constraint of HOG features is that it performs poorly when the background is noisy [5]. LPB is used to detect textures in order to overcome this shortcoming. LBP has been widely used in human detection and achieves good performance when combined with HOG [5].

Local binary patterns (LBP) is used to describe the texture of the images. LBP is extracted in a symmetric circular neighbourhood by comparing each pixel to its neighbour [5].



Here P is the number of neighbours and R is the radius of the neighborhood. Gi is the intensity of the neighbouring pixel and gc is the intensity of the center pixel [5]. u(x) is a step function where u(x) =1 when x>=0 and u(x) =0 otherwise. A histogram is created where there are bins for each uniform pattern. All non-uniform patterns are assigned to a single bin [5].

**Evaluation of Person Detector:**

**Describe how you evaluated the performance of your person detector. Take the ground truth to be the bounding boxes detected in part 1. Use the IoU metric to quantify the performance of the detector. Provide an image showing the detections obtained on at least one image from the dataset.**

**Evaluation of Person Counting:**

**Run the person detector of part 2 on all 1000 images of the dataset. Construct the response spreadsheet and upload it to the Kaggle competition web site. List the score (ranking and metric) reported on the Kaggle web site leaderboard. (note: this will not be the *final* score/ranking, as that will only be reported after the end of the competition period).**

**Discussion:**

**Discuss the lessons learned during this project. How could you improve upon your results? Are there better ways to estimate how many people are in an image than the approach used in this project?**

One of the shortcomings of LBP is that it cannot handle changes in illumination with much success. The results could be improved if a combination of HOG and LBPHF is used. This will greatly improve detection performance as can be seen in [5] where the HOG-LBPHF combination greatly outperforms HOG\_LBP.

**Describe any difficulties you faced in the project. Discuss the suitability of different types of features (e.g. HoG vs Haar or LBP or even Deep Features provided by a CNN).**

References

[1][Detectron2: A PyTorch-based modular object detection library](https://ai.facebook.com/blog/-detectron2-a-pytorch-based-modular-object-detection-library-/)

[2]<https://towardsdatascience.com/object-detection-in-6-steps-using-detectron2-705b92575578>

[3]S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017, doi: 10.1109/TPAMI.2016.2577031.

[4]M. Bertozzi, A. Broggi, M. Del Rose, M. Felisa, A. Rakotomamonjy and F. Suard, "A Pedestrian Detector Using Histograms of Oriented Gradients and a Support Vector Machine Classifier," 2007 IEEE Intelligent Transportation Systems Conference, Seattle, WA, 2007, pp. 143-148, doi: 10.1109/ITSC.2007.4357692.

[5]Aili Wang, Shiyu Dai, Mingji Yang and Y. Iwahori, "A novel human detection algorithm combining HOG with LBP histogram Fourier," 2015 10th International Conference on Communications and Networking in China (ChinaCom), Shanghai, 2015, pp. 793-797, doi: 10.1109/CHINACOM.2015.7498045.

[6]N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886-893 vol. 1, doi: 10.1109/CVPR.2005.177.