CITS4402 Computer Vision Report

Yao Tang 23702349
James Hudibyo 23452945
Tolu Daramola 23391695

May 12, 2025

1 Data Collection

The data comes from PTEA dataset and caltech-101 data set . Human and nonHuman data are selected from these set and preprocessed by resizing all images to $64 * 128(width * height)$. The human data are primarily come from the PTEA dataset, which provides images in 64*128 resolution by default.
The data set is organized into two folders: training set and testing set, each of these contains two subfolder: human and nonHuman, We use 800 images in total for training, 400 human and 400 nonHuman images. And 200 images for testing the result, 100 human and 100 nonHuman images.
The images are saved as JPG or PNG format. The images will be transformed to grayscale image only after reading.

2 Feature Extraction and Model Training

2.1 HOG settings

The default HOG settings used in the report are: cellSize = 8, blockSize = 2, bins = 9. There is a HOG class implemented, it will do three things in order: read and resize the image, compute HOG histogram and output the features.

2.2 SVM training

This report uses the sklearn library to execute SVM training. A linear kernel is used as required. The decision scores are calculated and used for plotting the MR vs FPPW curve. The training is run on all the images in the training set folder, with human images labeled as 1 and non-human images labeled as 0.

2.3 Evaluation Metric

Two metrics are used. The first is the accuracy score output by the prediction function of SVM. The second is the Miss Rate vs FPPW curve, which will be shown in detail in the Ablation Study section. The accuracy would be used as an indicator, a basic reference. While the MR vs FPPW curve is for further analysis.

3 Ablation Study

Four HOG settings are used for ablation study, which are concluded in the table below:

From the table, it is seen that except the 16*16 cell size configuration, the others achieve a 0.995 accuracy, which means only 1 image is wrongly identified. The MR vs FPPW curve is in the figure below:

Setting	Cell Size	Block Size	Bins	Accuracy
HOG-1	8×8	2×2	9	0.995
HOG-2	16×16	2×2	9	0.99
HOG-3	8×8	2×2	6	0.995
HOG-4	8×8	2×2	12	0.995

Table 1: Different HOG parameter settings used in the ablation study.

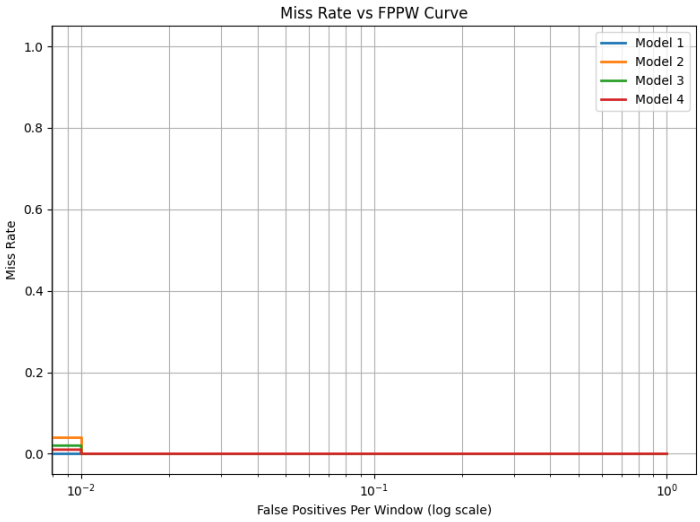

Figure 1: MR vs FPPW

From the curve we can see the result are nearly perfect, only tiny difference for different settings. According to our analysis, this is due to the relatively small size of sample space. From this experiment we can see smaller cell size tends to have better result.

References

[1] Dalal, N., Triggs, B. (2005). Histograms of oriented gradients for human detection. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (Vol. 1, pp. 886–893). IEEE. https://doi.org/10.1109/CVPR.2005.177