

1 Theory of the Processing

1.1 Smoothing

To help cope with noise, all my images went through an initial stage of smoothing. In the context of this problem, it was crucial that the edges of the pedestrian crossings remained well-defined and as sharp as possible. The reason for this is because I wanted my edge detection to work as good as possible. As a result, I settled with **median smoothing**, primarily because it does not blur edges too much.

1.2 Edge Detection

At first, I experimented with region processing using Mean Shift Segmentation (MSS). However, I found that it took about two seconds to process the image using MSS. In the context of the problem, I opted for a more efficient approach as it is vital that images are produced rapidly in a car imaging system. This is how I settled with edge processing.

I chose to go with **Canny Edge Detection** as opposed to Boundary Chain Codes, because with Canny I could have more customisation over the results by changing the parameters, whereas with Boundary Chain Codes it was more limiting.



Figure 1: Median Smoothing. I experimented with multiple iterations of median smoothing, however I found the best results with just one iteration. Additionally, I opted for a filter size of 3x3 as I wanted the edges to be as sharp as possible. With larger filter sizes, I was worried it would corrupt the image too much and deviate too far from the original.

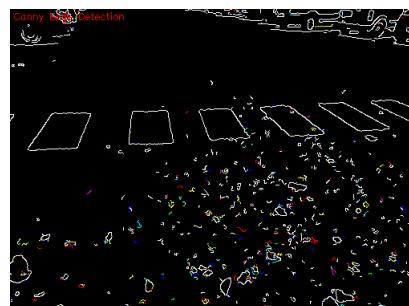


Figure 2: Canny Edge Detection. The low and high thresholds were chosen mostly based on trial-and-error. I ensured that my low threshold value was conservative, as this is indicative of the number of contours detected by Canny, and I wanted to be sure that my edge detection would not miss out on any potential pedestrian crossings.

1.3 Binary Thresholding

The output of the previous step in the process was an RGB multispectral image, and I wanted a binary image to simplify things further. I used **Otsu Thresholding** to convert to binary, followed by a closing operation. See figure 3 for results.

1.4 Connected Components Analysis and Shape Analysis

My next goal was to find contours in the binary image. I used CCA for its simplicity, and also because it showed great results on my initial testing.

After finding all the closed contours in the image, they would then go through the following filtering process:

Note: All parameters were chosen through trial-and-error testings.

1. Filter by Contour perimeter and area

- Contour Perimeter $>= 90$
- $50 <= \text{Contour Area} <= 3000$

2. Convex Hulls were then drawn around these contours. The Convex Hulls were filtered by area.

- $500 <= \text{Convex Hull Area} <= 3500$

3. Convex Hulls were then filtered by rectangularity.

- Rectangularity $>= 0.5$

After going through this filtering process, we now have a list of filtered convex hulls that potentially identify as pedestrian crossings.



Figure 3: Left-to-Right: Otsu Thresholding, Closing, CCA & Shape Analysis. The closing operation was done to fill any potential gaps in the edges of pedestrian crossings. This would yield better results for CCA. From some trial-and-error testing, I found a closing kernel size of 2 to be optimal for my process.

1.5 Colour Comparison

One distinctive trait of pedestrian crossings are it's white-ish colour. For each filtered convex hull, the mean colour of the convex hull is calculated. This RGB value is compared to the white RGB value by calculating the Euclidean distance between the two 3D vectors. From trial-and-error testing, I found 150.0 to be a forgiving value with good results.

1.6 Filtering out isolated & overlapping Convex Hulls

Another distinctive trait of pedestrian crossings is that they are positioned relatively close together. In this step of the process I wanted to filter out convex hulls that were isolated and had no neighbours - likely these are not pedestrian crossings because otherwise they would have neighbours.

This was achieved by finding the centre points of each convex hull and calculating its distance to the other convex hulls. If it had no convex hull within a distance of 175 (found through testing), it was filtered out.

Some convex hulls were overlapping with each other, so these were filtered out by introducing a minimum distance value of 10.



Figure 4: Left-to-right: Convex Hulls, White Regions, Filtering out isolated convex hulls. We can see the centre image has filtered out the van region in the top-right of the image because my the enclosed contour is not white enough. In the right image we can observe the leaves in the bottom right corner have been filtered out as they have no nearby neighbours. If you look closely at the pedestrian crossings, we can observe that overlapping edges have also been filtered out.

1.7 Finding the longest linear sequence

A key trait of pedestrian crossings is that they occur in a straight line. The aim of this step was to determine the longest linear sequence among the potential pedestrian crossings. This was achieved with a brute-force approach - in reality, there will only be a handful of potential pedestrian crossings so any efficiency improvements would be negligible.

A line is drawn from the centre of one convex hull to a second convex hull. Then, another line is drawn from the second convex hull to a third convex hull. If the angle between these two lines is less than 8° , it is deemed in a straight line and the algorithm continues to find other convex hulls that also lie within this threshold. This is repeated for every possible combination of convex hulls, and then the longest linear sequence is obtained.

Additional Heuristics

- If there is a tie for the longest sequence, my algorithm will choose the linear sequence that is 'straighter'.
- A linear sequence must be of length ≥ 3 to be considered a pedestrian crossing.
- My algorithm will only return one potential pedestrian crossing. If there are multiple pedestrian crossings in the image, it will choose the longer one.

1.8 Drawing the Bounding Box

If a pedestrian crossing was identified, a bounding box was drawn around it.

Various straight lines were drawn through the centres of the pedestrian crossings and the mean slope was calculated. The top edge of the bounding box would be drawn through the maximum y-coordinate with this mean slope, and the bottom edge of the bounding box would be drawn through the minimum y-coordinate with this mean slope.



Figure 5: Longest Linear Sequence.



Figure 6: Drawing the bounding box.

2 Results and Analysis

2.1 Performance Metrics

Seeing as this is an object detection problem, it made sense to use Intersection over Union (IoU) and the F1 Score as my performance metrics.

$$IoU = \frac{Predicted \cap Ground\ Truth}{Predicted \cup Ground\ Truth} = \frac{Intersection\ Area}{Ground\ Truth\ Area + Predicted\ Area - Intersection\ Area}$$

The *Intersection Area* was calculated using the `intersectConvexConvex` method. The *Ground Truth Area* and *Predicted Area* were calculated using the `contourArea` method.

Additionally, the F1 score will be used in measuring my algorithm's performance based on both training and test data. Following the advice from the lectures, a True Positive (TP) is when the IoU is over 50%. I chose to go with this metric as it integrates elements of both Precision and Recall.

$$F1\ Score = \frac{2 * TP}{2 * TP + FP + FN}$$

For the 20 training & test images, the F1 Score of my solution was 100%, meaning it correctly identified all True Positives and True Negatives with no mistakes. The mean IoU across the True Positives was 84.76%; the lowest value being 72%. In my view, this is a strong result seeing as an IoU over 50% is considered to be an observed positive.

2.2 Presenting Visual Results

I will not be including all cases of successes (as many of them are trivial), only examples of significance. The green lines correspond to the Ground Truth, and blue lines correspond to my Prediction.



Figure 7: We can observe that even if all individual crossings are not captured, the bounding box gets extended to capture the entire crossing, demonstrating the robustness of the solution. The limitations of this will be discussed in figure 12. IoU=78%.



Figure 8: Despite some worn-down crossings, all are identified. The colour distance threshold had to be tweaked to a forgiving value (150) to consider edge cases such as the above. This also has its limitations and will be discussed in section 2.4.2. IoU=84%.

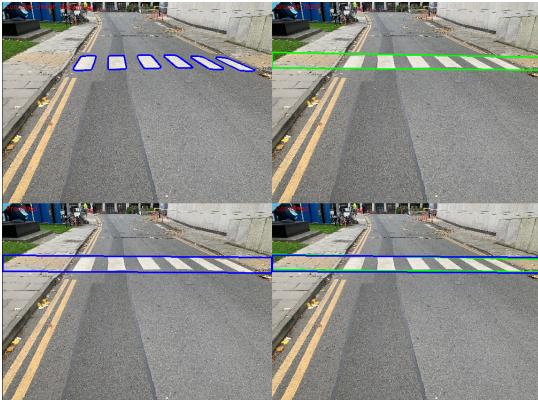


Figure 9: We can observe that the solution still identifies crossings far away at a small size. The minimum area threshold had to be tweaked carefully to allow for this. IoU=88%.

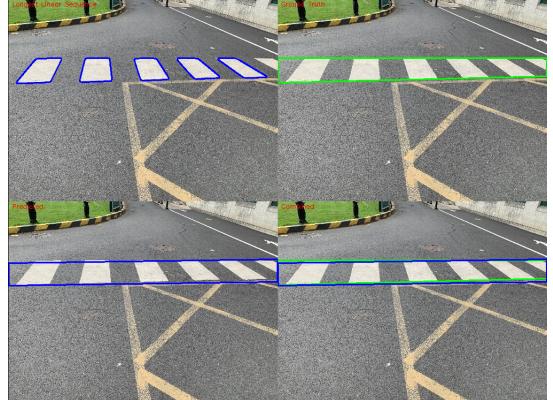


Figure 10: Despite some other yellow road markings near the crossing, it does not affect the results of my solution. This is a result of resilient colour distancing and shape analysis. IoU=88%.

2.3 Robustness

2.3.1 Spacing

A key feature of pedestrian crossings is that they are relatively evenly spaced. I was considering integrating this into my solution by examining the relative distance between each of the identified crossings, but I decided against it for two main reasons:

1. The distance between the crossings would vary drastically due to perspective warping. This would be a very challenging obstacle to overcome.
2. If one of the crossings in the middle would fail to be identified - for example, it could be covered in leaves - it would collapse my solution and fail to recognise the entire crossing. See figure 11 for a good visual example.

2.3.2 Outlier Resilience

As mentioned in section 1.7 & 1.8, when checking for collinearity between the crossings, I allow for 8° deviation. Furthermore, the gradient of the bounding box is the mean slope between all the centroids of the identified crossings. This means that even when there is an outlier crossing, it will not significantly hurt the output of the solution. This is best demonstrated in Figure 12.



Figure 11: Despite the solution failing to identify the crossing in the middle, it still recognises the other ones and since they remain in a linear sequence, it correctly identifies it as a pedestrian crossing. IoU=72%.

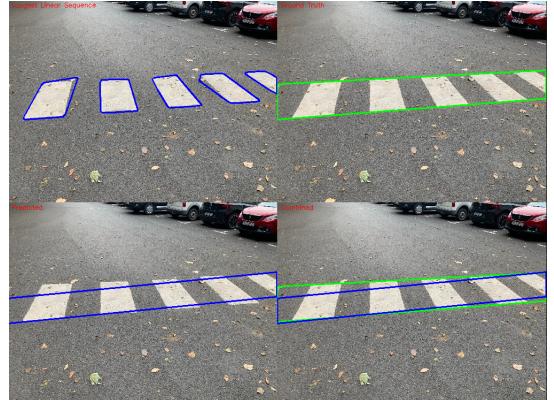


Figure 12: Since the crossing on the right side is cut off, the centroid of that crossing skews the predicted bounding box. Despite this, my solution still works well in identifying the location of the crossing with an IoU of 77%. While it's not perfect, it still performs well above 50%.

2.3.3 Non-Pedestrian Crossing Images

I also wanted to consider cases where the images do not feature pedestrian crossings. Road markings were an obvious edge case - the key feature about those is that they run relatively vertical to the perspective of the camera, whereas pedestrian crossings run relatively horizontal to the camera.

To counter this, I integrated a feature that compares the angle of the identified pedestrian crossing to a horizontal line - if it lies under 30°, then it is considered a pedestrian crossing.

Otherwise, it is likely a road marking. This value was obtained from various trial-and-error testing.



Figure 13: Despite identifying some potential candidates for crossings, they are rejected as they are deemed 'too vertical'.



Figure 14: The colour distancing does most of the heavy lifting on this one, but even if the markings were perfectly white, the horizontalness check would have rejected it.

2.4 Limitations & Potential Solutions

In this section, I will discuss my solution's limitations and their potential solutions.

2.4.1 Shape Analysis

In section 1.4 I mention that my threshold value for rectangularity was 0.5. Intuitively, this sounds like a very loose value. However, as I tested my solution on the training images, the results were very poor for high values of rectangularity, so I lowered the value. See figure 15.

As a potential solution to this, I could have performed a perspective transformation on the individual crossing, and then perform the rectangularity analysis. This would result in more realistic and accurate shape analysis.

2.4.2 Colour Distance

In section 1.5 my method for comparing enclosed contours with white is described. Despite my best efforts for calculating the most optimal colour distance value, the results were shaky at times. See figure 16 for an example. Furthermore, I doubt my colour distance method would work well in darker lighting conditions.

This could potentially be improved by obtaining a dataset of varying pedestrian crossings and performing histogram back projection to obtain a probability image. We can threshold for values above 0.8 for example and use these as our white regions.

2.4.3 Reliance on 'Magic Numbers'

My solution features 13 constants that were fine-tuned by hand. I tried my best to make the values loose and forgiving, but the reality is that my reference was only the 10 training images. As a result, it is likely that my parameters were over-fitting to the training set. Normally an indication of over-fitting is poor performance on unseen test data. In my case, my solution worked very well on the unseen images.

However, I would argue that the unseen image test set was very generous - the photos were mostly of the same pedestrian crossings at different distances. I doubt its performance would be as good on another diverse image set.

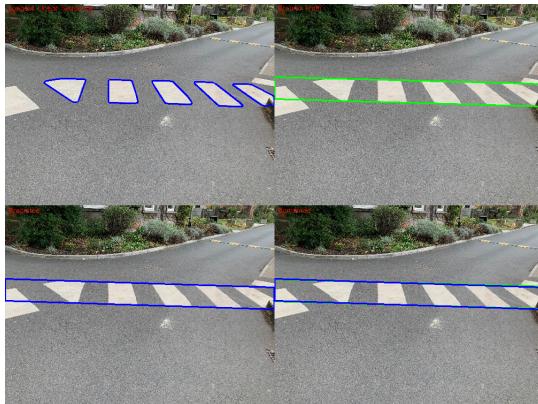


Figure 15: The triangular pedestrian crossing luckily gets identified. However, I would rather this wasn't the case - from an intuitive point-of-view, this is not a rectangular shape, so it should not really be identified as a crossing.

2.5 Overall Results

Image	IoU	TP	FP	FN	TN
PC10	87%	✓			
PC11	73%	✓			
PC12	90%	✓			
PC13	88%	✓			
PC14	88%	✓			
PC15	86%	✓			
PC16	78%	✓			
PC17	89%	✓			
PC18	84%	✓			
PC19	N/A				✓
PC20	77%	✓			
PC21	93%	✓			
PC22	85%	✓			
PC23	94%	✓			
PC24	72%	✓			
PC25	75%	✓			
PC26	87%	✓			
PC27	87%	✓			
PC28	N/A				✓
PC29	N/A				✓



Figure 16: My solution still recognises two letters as potential pedestrian crossings. Recall a heuristic I implemented was that for something to be recognised as a pedestrian crossing, ≥ 3 individual pedestrian crossings must be identified.