IVR Coursework 1
Report

Weiting Goh (S1450710) and Tomas Markevicius(S1452595)

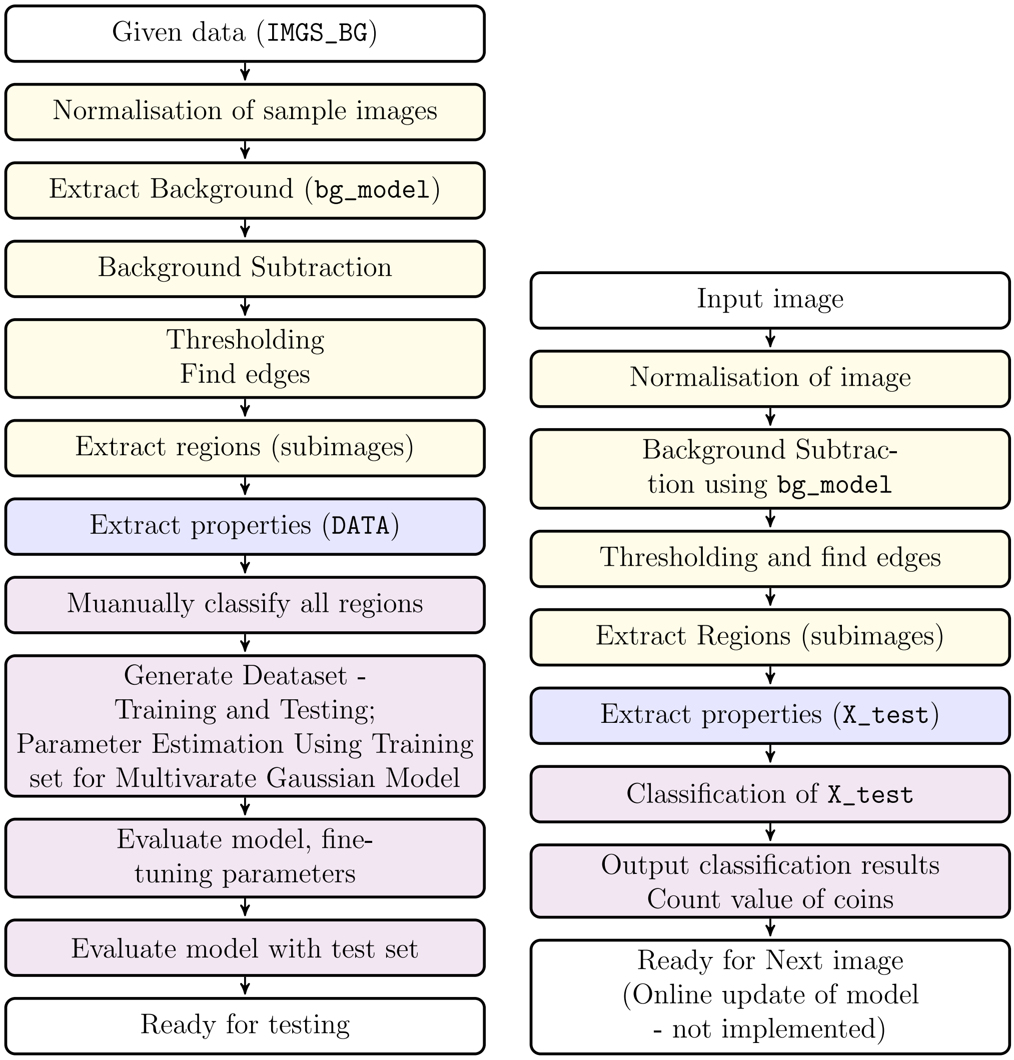
October 26, 2016

1 Introduction

To recognise and count the coins in an image, Coinsy, have three subtasks: 1) Image processing (includes image segmentation and smoothing), 2) Feature Extraction, and 3) Classifiation (and then counting the coins).

For Coinsy to be a proficient counter, we have to first train it. The processes that Coinsy goes through for training differs from evaluation as described in figure Figure 1. We will describe her training in the next section - methodology, and her evaluation results after. Lastly, we conclude with a discussion on Coinsy performance.

In the following subsections, we give an overview of the operation pipeline for each subtasks. The approach here is an abstract idea of what we did for our base model. Additional techniques were explored and will be discussed in later section.



(a) Operation pipeline to train Coinsy

(b) Pipeline for (trained) Coinsy to count coins

Figure 1: Differences in Pipeline; Yellow boxes indicates the image processing procedures; Blue indicates the feature extraction procedures; Violet indicates the classification procedures.

1.1 Data

The data we were given are images (split into two dataset - `harder` and `simpler`) with objects in the foreground for us to classify. The objects differ in colors and some are very similar to the background - such as 1 pound coins (see Figure 2). Hence, the classifier must be invariant to rotation of objects, and importantly to detect the objects will the images to be processed, such that the objects are salient to the computer vision. The real challenge lies in segmentation of the image.

Figure 2: Sample images from given data.

Figure 3: Sample images from given data.

1.2 Image Processing

The image processing step aims to 1) make all images (for training the model or for evaluation) comparable, 2) extract objects in the foreground (also known as image segmentation). The outcome of this stage is an array of subimages ready for feature extraction.

The central idea is to, first, model the background before subtracting it from all images; second make the edges of the objects 'obvious' by thresholding; third, crop out the objects to obtain the subimages. This is one of the many image segmentation technique available, and further exploration of other techniques is discussed later.

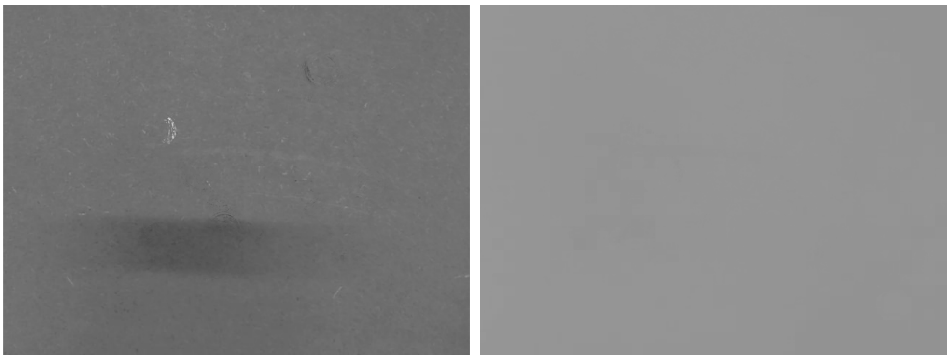
Background Subtraction

In this coursework, since a background image is not readily available, we have to model it. Noting that images varies in illumination, we have to make the images comparable by normalising it first, using the following formula.

$$P_{r,c}(R', G', B') = (\frac{R}{\sqrt(R^2 + G^2 + B^2)}, \frac{G}{\sqrt(R^2 + G^2 + B^2)}, \frac{B}{\sqrt(R^2 + G^2 + B^2)})$$

With objects scattered around randomly in the images, we find the median of all image pixels for each channel separately in order to reconstruct the background.

The outcome of the background with and without normalisation is shown in Figure 4.



(a) Median Filtering without normalisation (b) Median Filtering with normalisation

Figure 4: Background model generated from all 14 images

The sample images in Figure 2 after their background removed is shown in Figure 3. It is evident that the background removal process renderede a large part of the background

Segmentation

1.3 Classification

2 Methodology

THIS IS THE TARGET

3 Result

Appendix here