

Object Detection and Counting Based on Modern Computer Vision Methods

Wenjin Yang, Linlin Ouyang, Qingting Jiang, Wuchen Li, Tianran Zhang
ShanghaiTech University
{yangwj1, ouyangll, jiangqt, liwch, zhangtr}@shanghaitech.edu.cn



Goal

For this project, our goal is to detect all objects in a scene and count them indiscriminately in order to get the degree of chaos in the current scene.

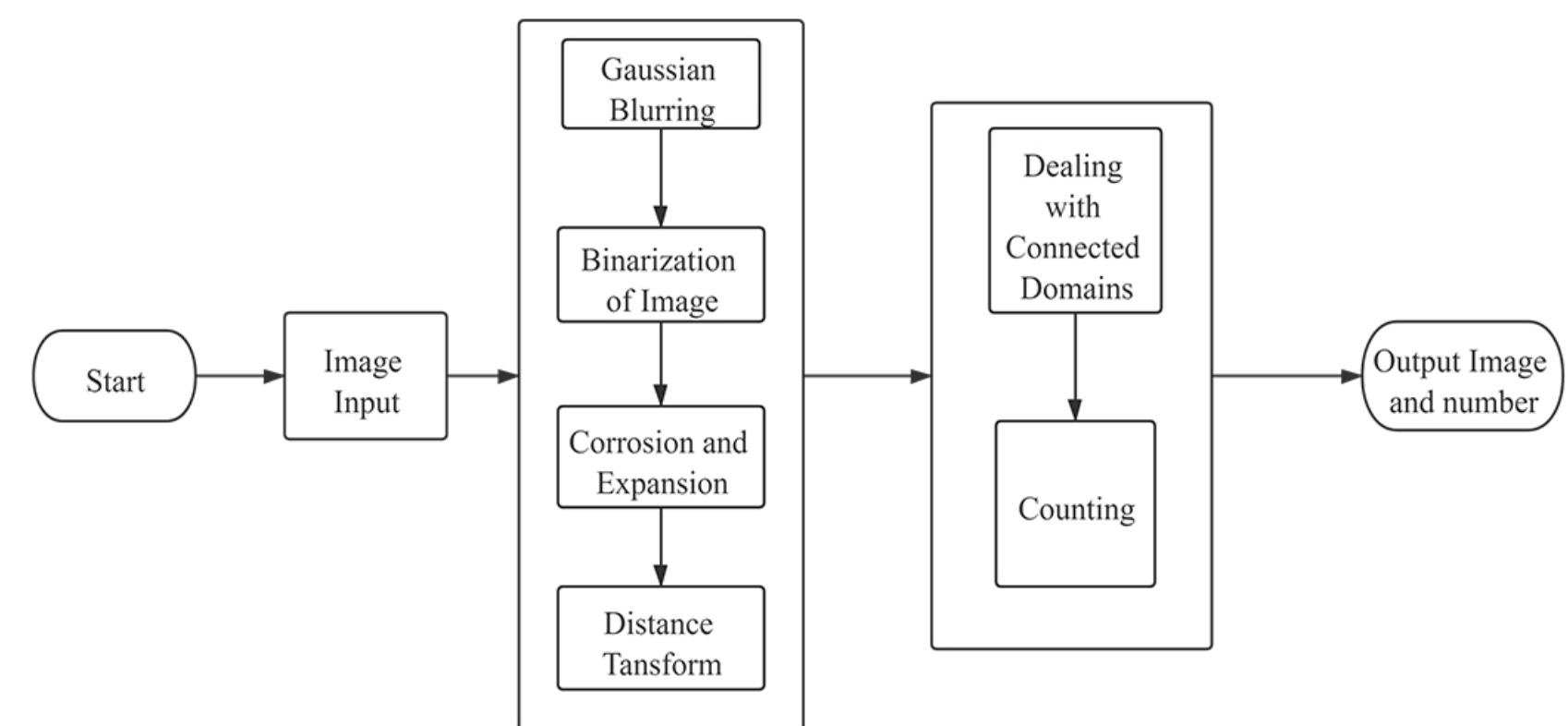


Contribution

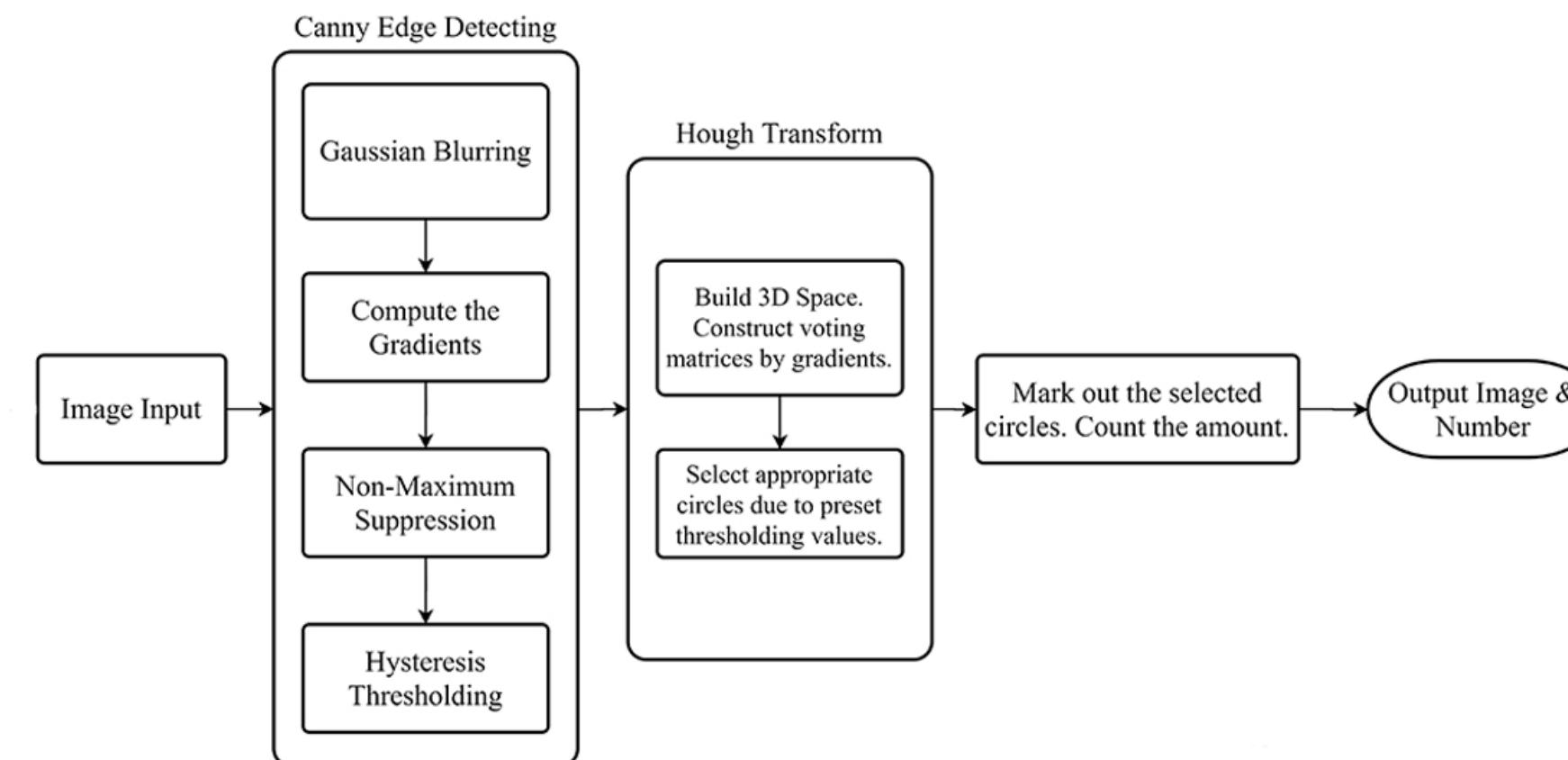
Our algorithm can count all the objects in a certain scene and get the relative size of each object (we use the relative size instead of the absolute size because we expand and corrode the image here, and the relative size has no effect on our judgment of the degree of chaos). This technology can use the relative size of each object to calculate entropy to detect the degree of chaos in the scene.

Methods

1.Watershed Algorithm: Watershed segmentation is a natural-based heuristic algorithm, the basic principle is to regard image features as geomorphological features, and to divide each area by using the grayscale distribution characteristics of pixels. Through the watershed algorithm, we can convert objects in the image into a closed area, then by calculating the number of connected domains, we get the number of objects.

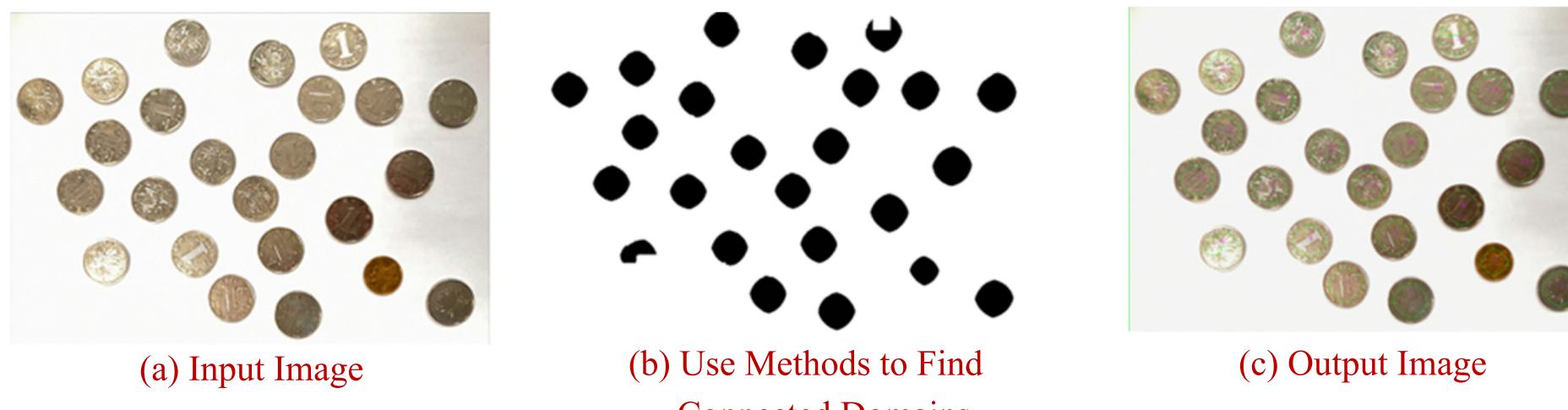


2.Canny Edge Detection & Hough Transform: In Canny, using a Gaussian kernel can filter out noises and Hysteresis thresholding to decide whether a weak edge should be adopted. Hough Transform builds a 3D space by x, y and radius, and calculate a vote matrix according to the gradients. Thus circles can be selected from the matrix limited by thresholds.



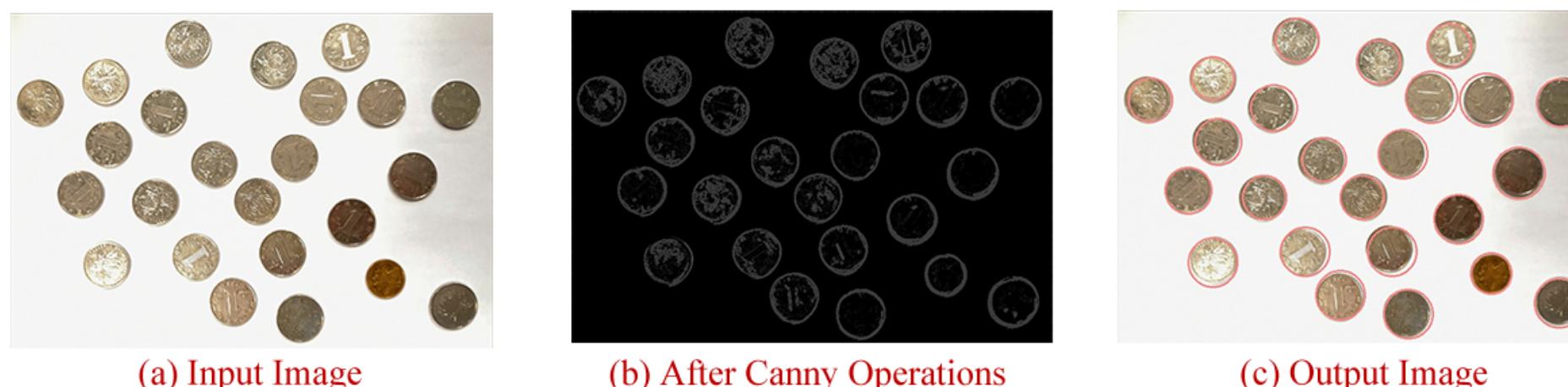
Experiment Results & Analysis

1.Results of Watershed Algorithm:



Error Analysis: Its noise-sensitivity would bring massive errors, while It is easy to produce over segmentation .Also, it is greatly affected by the change of image light source. Important contours are easy to lose for low-contrast images.

2.Results of Canny Edge Detection and Hough Transform:



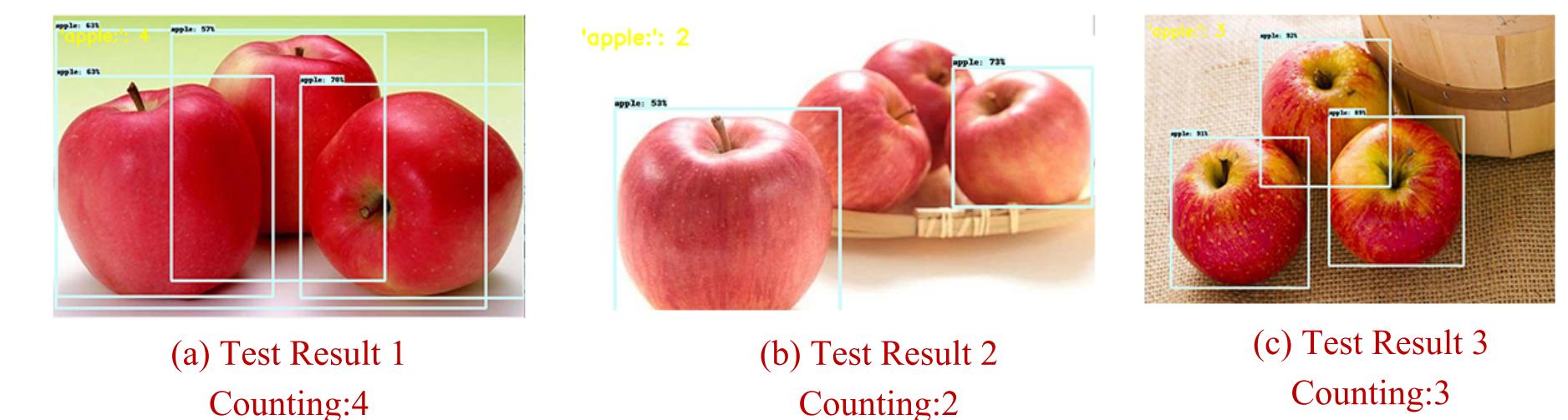
Error Analysis: Since the thresholds in Hysteresis Thresholding decides the range of weak edges, some edges may not be detected under high thresholds. Also, a low threshold when selecting circles in Hough Transform can lead to a multiple detection within one circle.

Advanced Model

Theory: This model is implemented base on SSD-Mobilenet model, serving as a testing benchmark and a new approach to solving image detection issues with object occlusion. Feature extracting is based on trained color histogram features and pixel gradient distance. Model is pretrained with COCOv1 dataset and

achieve a good performance in object detection and counting. Object classification and object counting is two task implement by two part of object detection but during the model training the features used to classify object is also used to clarify the boundary of a object. Object counting on specified object training is ought to be explored in the future.

Test Results:



Conclusion

Watershed Algorithm

Pros:

- 1.It can detect a large number of objects accurately.
- 2.It can detect objects of all shapes and types.

Cons:

- 1.Excessive segmentation, An object may be divided into multiple connected domains due to details such as the pattern of the object itself.
- 2.Highlights on objects can affect detection results.
- 3.It is not robust to the situation in which an object is obscured.

Canny Edge Detection and Hough Transform

Pros:

- 1.It is a simple and basic way to find out circle edges.
- 2.Normally, it can get the result in a short time.

Cons:

- 1.It cannot correctly count object with overlapping.
- 2.A little change in parameters may lead to a different outcome.
- 3.Different inputs have different best parameters.

Advanced Model

Pros:

In object detection and counting, it is one of the most efficient methods among modern computer vision methods, which could usually reach high performance.

Cons:

Even with varied-size of features extracted by SSD, size changing has a highly influence on object detection. The model used is trained by a general dataset with multiple types of objects so dataset may not be large enough to achieve a high performance on apple counting.