Ordered Object Segmentation Based on Machine Learning Algorithms

Md Arif Uddin[*1], Hongbin Ma[*2], Haotian Wu[*3], and Ying Jin[*4]
Beijing Institute of Technology Beijing 100081, China
E-mail: [*1] arifuddin0809@gmail.com , [*2] mathmhb@bit.edu.cn
[*3] wht_bit@163.com, [*4] jinyinghappy@bit.edu.cn

Abstract: The conventional automatic drug dispenser can significantly contribute to numerous services such as medicine industries, pharmacy shops and medical assistance by reducing the workload of medical workers and allowing patients to take their timely medication. In this paper, we use a segmentation technique to identify the location of drug pillboxes from the medicine shelves with the automatic counting of spiral pillboxes. The proposed method is a solution to a dynamic segmentation framework for recognizing objects from large datasets. Firstly, the picture is filtered to remove the noise to recognize the boundary line. Following that, projection lines are drawn in the center to divide the segments of pillboxes by utilizing the midpoint formula, which supports smooth counting of line detection among objects. Besides, we applied the Mean shift method to accumulate the intersection points. The key role of clustering is to collect intersection midpoints data. The proposed framework identifies the existing area from medicine pillbox images and detects the specific segmentation line that automatically allocates the pillboxes. Furthermore, we obtain 99.83% accuracy and successfully detect pillboxes, which aids in the improvement of segmentation, filtering and clustering algorithms. Such techniques assist us in recognizing the most appropriate segmentation boundary lines for further discoveries in the future.

Keywords: Hough lines, detection segmentation lines, intersection lines, mean-shift cluster, intersection points cluster.

1. INTRODUCTION

At present, image processing is frequently used in several studies to extract various types of images, such as real-time image processing, identifying objects, face recognition and human handwriting recognition [1, 2].

Recently, the medical industry, healthcare industry and pharmaceutical industry companies in this area have seen significant errors and mistakes in calculating the data. They are having difficulty in estimating the missing things, the procedure is also getting increasingly laborious, time-consuming and such counting is a difficult activity for humans.In this case, it can be solved by using artificial intelligence through a robot to count automatically and give the solution without making any mistakes, for example, capture image, call number detection, detect the position, detect the title, RFID tag also scan the bar-code so on. Many recent works have produced algorithms for extracting book titles and individually identifying and recognizing books from shelves or identification, detection and pulling books from shelves [3]. In this paper, we focused on segmentation and counting medicine pillboxes and books, which is the main aim of this research. It is a great challenge in this research area.

The primary contributions of this work are as follows:

- The image processing technique is applied to use the medicine pillbox images in different shapes.
- We use a method to identify the space between the pillboxes and draw segmentation lines.
- The empty space between the objects on the shelves is detected and shown as the missing object. Also can tell us about the boxes which are not in the vertical position, resulting in various angle forms. Despite some pictures being dark, others being bright.
- We sketch projection lines in the middle to divide the segments of pillboxes. The method helps to count and detect lines among objects smoothly at the intersection points.
- Circles are drawn on each of the clustering intersection points to compute all the clustering points.

The remainder of the paper is organized as follows. In Section 2, we highlight the existing works and highlight the novelty of our work. Section 3 illustrates the proposed method in detail. The experimental process and results are discussed in Section 4. Finally, we conclude the paper in Section 5.

2. RELATED WORK

Many researchers have presented several types of technologies. There are several studies on book detection and segmentation where authors have proposed different ways to get a book out of bookshelf images. In [4, 5] consider different shapes angular on a shelf in exact lines and detect individual books. Based on boundary detection, [6] has only taken the book boundary detection technique where the authors used horizontal projection to find the boundaries. Besides, the librarian robot is described in [7], as well as automatic book recognition through optical character recognition, which is visually guided. the image processing system used in counting automatic books in a library creates a ROI from the input picture of the book's current location [8]. Further, the Canny edge detection algorithm is then used to extract book edges. We use similar logic in this work. However, we trace the intersection line in the center of the segmented lines while also observe the noise in the boundary lines. At the same time, we identify

the edges and determine the locations of the spaces. Here, [9] the Canny edge detector produces curved or broken lines after recognition of the call number label on the book spine captured by a smartphone and book spine segmentation. To group linked components of characters on the edge map, use Contour clustering. To discuss a method for segmenting books spine and recognizing the title [10]. Specific rows and characters of the title are segmented and extracted using a bounding box and linked part field. Besides, [11] is an interesting research method that indicated Haar and detection of Ada Boosts call number which further indicated training sample data at 1000 only. The proposed clustering method uses GPS and road intersection point traces [12]. The authors use the Mean-shift algorithm to raise the maximum density of neighborhood points and to improve the geometric accuracy of intersections is to identify the dominant orientations of road segments. The mean-shift algorithm is an effective method for cluster neighbor used multi-manifolds intersection cluster uses segmentation, the nearest point cluster, and includes the tackling geometric intersection points cluster [13].

Although most literature focuses on the image segmentation method but they did not apply the clustering method to perform book detection. Therefore, we use search results of the current method medicine pillboxes and books to detect the ability of segmentation from the image of medicine pillbox shelves. The technique works through numerous case studies in the pillbox-shelf remedies and book stories.

3. PROPOSED INTERSECTION POINT CLUSTERING TECHNIQUE

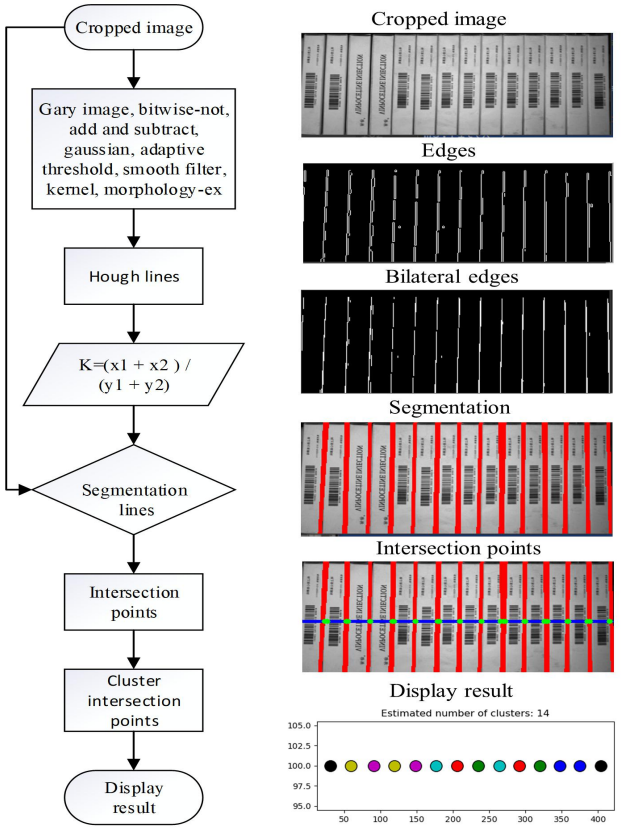

Fig. 1 The extraction method of segmentation lines and cluster intersect points.

In this section, we discuss the proposed clustering technique and its major steps in detail. The proposed intersection point clustering technique is divided into five major steps: 1) Extraction of the Boundary Lines 2) Extract Segmentation Lines and Intersection Points 3) Mean-Shift 4) Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and 5) Center-Cluster and Counting Boxes. **Fig. 1** shows the principal design platform step by step.

3.1. Extraction of the Boundary Lines

3.1.1. Pre-process

After Grayscale the Bit wise-Not operation is processed which gives a binary image. This is a great procedure to detect binary objects. Whether the object is black or white, we use the addition or subtraction method depending on the image's object type. In the next step, we used dark and light pictures and tested them in this field. Furthermore, the Gaussian blur technique is commonly used to reduce image noise and information. The adaptive threshold is a process that determines the value for smaller regions, resulting in varying threshold values for different areas. It detects the boundary lines smoothly and adds them to the kernel for small matrices which is one. It was a working image in black and white and took on to the numbers from 0 to 255 pixels detected in the corresponding brightness pixel image. Finally, morphology is a function applied to preset pixels that compare the output image with the input. Morphology works for two senses, always used in pixels to dilate and erode. Dilate is add the boundary lines, hence erode is the opposite to dilate. Initially, dilate was followed by eroded and the eroded was followed by dilatation. The structural image in **Fig. 2** shows the boundary for vertical or horizontal segmentation to extract the objects corresponding to their elements. The image segmentation depends on the picture being vertical and horizontal on the pillbox or bookshelf.

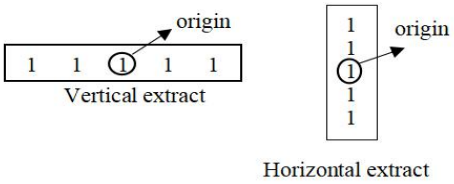

Fig. 2 [14] Extraction of the vertical and horizontal lines.

3.1.2. Possible boundary

This study focuses on one of the most important aspects of boundary extraction and identifies possible boundaries. Additionally, our project used books and images of medicine pillboxes after filtering for canny edge detection. Boundary recognition is the primary problem of computer vision. It isolates by the canny edge, which invents the boundary between brightness and dark pixel in an image. At the same time, Hough transforms to illustrate the straight line.

$$y = mx + b \quad (1)$$

We are familiar with the "$y = mx + b$" form (also known as the slope-intercept form of a line equation). To get the vertical line, we use the Hough line. The straight line depicts the position of (m, b) parameters as points.

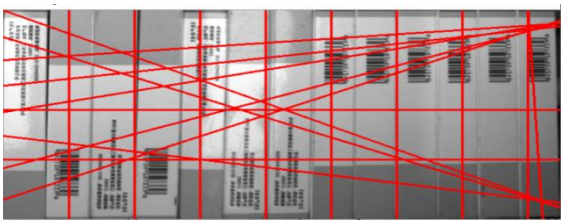

Fig. 3 Possible boundary lines.

$$r = x\cos\theta + y\sin\theta \quad (2)$$

The distance from the origin is r, and the angle between the x-axis and the closed points is θ. As a result, each line of the picture pair should be identified with (r, θ).

$$k = (x_1 + x_2)/(y_1 + y_2) \quad (3)$$

Equation (3) determine the slope k.

$$x_1 = \text{int}\,(x_0 + 1000 * (-b))$$
$$y_1 = \text{int}\,(y_0 + 1000 * (a))$$
$$x_2 = \text{int}\,(x_0 - 1000 * (-b))$$
$$y_2 = \text{int}\,(y_0 - 1000 * (a))$$
$$k = (x_1 + x_2)/(y_1 + y_2)$$

Fig. 4 After filtering, the vertical lines were extracted.

Here, we manipulate the parameter of k to cause horizontal lines to be more than the vertical lines. We then change the values and adjust the limit lines for vertical segmentation.

3.2. Extraction segmentation and intersection points

In this section, we discuss the extraction segmentation and intersection points method. We start by explaining the filtering technique and then elaborate on the drug box experimentation.

3.2.1. Filtering the boundary

Open CV conveniently draws segment lines using the Hough line transformation method to detect boundary lines [15, 16]. At this point, in **Fig. 5** we compute the result of the segmentation lines to obtain the boundary line here. It is important to note that the process faces the challenge of the segmentation boundary drawing numerous lines instead of a single line. As a result, of the solution of more lines, the modified Hough line threshold parameter has increased, which support us to decrease lines. whereas more than one line is not feasible for counting pillboxes, so the solution for this is that, we draw intersection lines in the center to divide the segments of drug boxes. In the future, the experiment will better way come to solve this kind of problem.

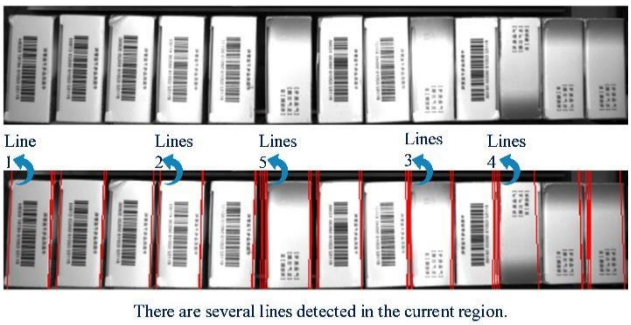

Fig. 5 Segmentation lines

$$\text{Midpoint} = \text{int}\,(x_1 + x_2)\,/\,2$$

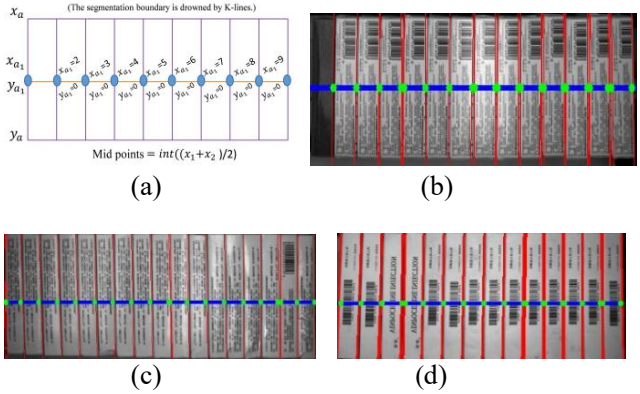

Fig. 6 (a) Midpoint formula and intersecting lines, (b) Dark image intersection points (c) Gray and dark image intersection points (d) Intersection points over bar code image.

According to the image, **Fig. 5** we used a method to solve the experiment of drug pillboxes and books counting. It is taken with the assistance of an intersecting line. In the following step, drawn the segmentation lines to use the Hough line technique. The projection lines are drawn in the middle of **Fig. 6** to divide the segments of pillboxes which assist to find intersecting points from each segmentation of the boundary line. The mid-points of the intersection points are calculated using the midpoint formula.

3.3. Mean-shift cluster

The Mean shift algorithm is used for image processing as the segmentation and the most commonly used centroid [17]. It is essentially the working model of a density function using kernel input. However, Mean shift is also known as kernel density estimation (KDE). In general, density estimation uses different points to make a group of multiple clusters. Here, the kernel weight function averagely assigns each data point. It usually follows as a distribution feature where it refers to the Gaussian distribution and provides bandwidth to KDE. For example, if the bandwidth is greater than 2, it will build clips or climbing a hill on the KDE surface to the nearest peak. It achieves such outcomes by slowly moving each point upward until it reaches a peak.

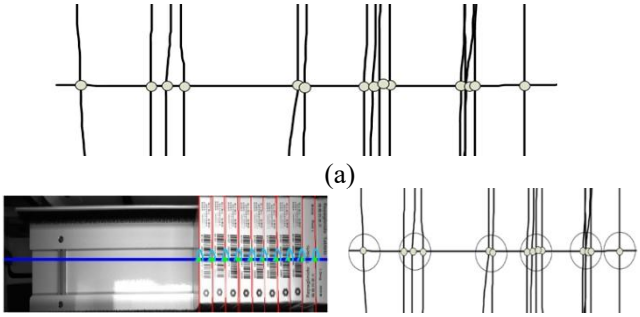

Fig. 7 Intersected points, clustering detection (a) Intersection points (Sketch), (b) Clustered points (Pillboxes) and (c) Clustered points (Sketch).

We adopt the concept from [18, 19] to work with discrete data convergence using a segmentation image, and then we use the mean shift algorithm. **Fig. 7** we obtain all the data points at the intersection points after the segmentation process. Besides, we achieve many lines beyond the boundary line. In **Fig. 7 (a)** on the demarcation line, there are multiple intersection points, which complicates counting the segmentation lines. However, we use the mean shift algorithm to solve this problem. We notice more than one line in the middle when plotting a line in the center of a segmented line. Therefore, in **Fig. 7 (b)** we use the clustering approach to make one point for each segmentation line. **Fig. 7 (c)** This technique helps to cluster nearby points by requiring one point on a different segmentation line and forms different color of circles in each segmentation line. The mean shift approach is used to identify regions in the data set where there are a lot of data points. The proposal of a staged hybrid clustering approach produces a high average outcome for every point cluster. However, the clustering produces a decent accuracy result for each segmentation boundary.

3.3.1. Mean shift cluster finds cross-pointing points

First, the group of intersection points based on the segmentation lines feature label datasets, bandwidth, quantile and sample estimate, etc. After using the mean shift cluster algorithm, then fit the data and labels, in addition at the same time clustering the intersection points. We learned before finding intersection points by using the Basis formula. An Intersection is that two sets of elements are used in the intersection point algorithm in mathematics. To intersect is two sets of elements A and B, here A belongs to B and B belongs to A, so, algorithm improved the clustering process divided into two steps separating and merging. [20, 21] the improved SC-KH (Staged text clustering) algorithm and the KM algorithm outperform both K-means and hierarchical clustering in terms of efficiency. On the other hand, we use the hierarchical approach to find the closest points. In this method of tracing the nearest neighbor obtains the date and a pair of the number of the (n) data which has various advantages.

$$A \cap B = \{x | x \in A \cap x \in B\}\ A \cap B = \varphi \qquad (5)$$

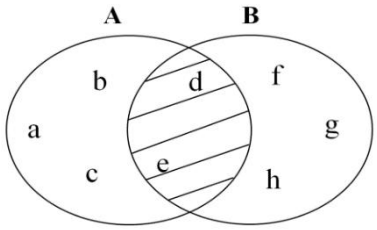

Fig. 8 Example of a figure caption.

Here, [22] hierarchical cluster algorithm also calls (NN), which means the nearest neighbor finds close neighbor points and make a group of clustering. The algorithm uses a pair of the nearest numbers of (n) data points and pairs them on the intersection points and also uses the measure of the distance between points to make (n × n) distance matrices and row data. It automatically clusters each other to find the nearest cluster together and then merges the similar cluster.

$$NN = \min_v \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \qquad (6)$$

Here, min is neighbor (v) data points all are taken and each
of the data points makes a pair. Index data points I_0p and their NNI_0NN will be used for making pairs.
Pairs = $[\{I_0p_1, I_0NN_1\}\{I_0p_2\ ,\ I_0NN_2\}...\{I_0p_n, I_0NN_n\}]$
Moreover, I_0p_1 first data set of points and index data points 1st data set I_0NN_1 is the NN first data point.

3.4. DBSCAN

Use unbalance data, [23] DBSCAN cluster algorithm set each grid and adjust to cluster mixed data zones. DBSCAN has several kinds of applications that have analyzed the data set but most of the time used data mining. For cluster commonly used DBSCAN and mean-shift clustering, DBSCAN is one of the famous algorithms, where it identified the different shapes that defined the number of clustering. It is commonly used data mining, an association of the points and the groups together make points and cluster of the nearest points with each other according to the measurements and a minimum number of dot and low-density regions. For this purpose, in addition to DBSCAN, we used the mean shift algorithm in our analysis. In our experiment, we have to test these two algorithms to see which one is the stronger cluster. Then we'll remember, and it'll be possible to consider which one to use in the future. Here **Fig. 9** used the DBSCAN method to set up the intersection data set points for each of the segmentation branches, then cluster the intersection point. This picture shows the effects of the books shelf cluster image. There are no noises in this image.

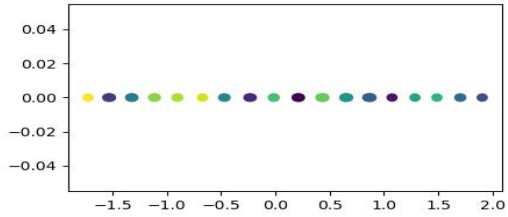

Fig. 9 Cluster of books shelf image from DBSCAN

Table. 1 Comparison of cluster intersection points.

Name	Segmentation image clustering				
	album	neighbor	accuracy	cluster	noise
Mean shift	1020	100%	98%	99%	no noise
DBSCAN	1020	98%	94%	97%	noise

Table 1 compares the two best clusters frequently used in most of the research. Here the clustering role is, to cluster the intersection points and estimate the number of the objects. This paper compares the clustering efficiency and accuracy where the mean-shift cluster neighbor points are caught up to 100% and DBSCAN 98%, accuracy 94%, cluster 97%, and noises. Mean shift cluster accuracy 98%, cluster 99%, and no noise.

3.5. Draw center-cluster and counting boxes

3.6.1. Segmentation and counting

We developed the circle on the intersection points as the final move after clustering. There are much more lines to identify between the drug boxes segmentation lines, which is difficult to estimate because of the enormous number of datasets. However, more than one line is inconvenient to count. The solution is to use the intersection process, to draw a midpoint line for every segment's branch over the segmentation lines to get the intersection points. After applying the clustering technique, we accumulated each of the intersection points. As a result, we clustered the intersection lines and estimated the clustered points using the circle center clustering process. The effect of this approach is more precise when counting medicine pillboxes and books. Therefore, no need to count separately on a large scale because through this method, we automatically obtain all the segmentation lines.

4. EXPERIMENTS AND RESULTS ANALYSIS

The proposed approach was tested in an OpenCV environment using a core-i7 2.59GHz processor and 16 GB of RAM. For this technique, several types of drug images are detected from the medicine shelves also book stores and applied as input data. The result of applying the desired method to the different medication pillbox shelves photos is shown in the simulation figures. The first figure in **Fig. 10 (a**) describes the number of clustering (20 pillboxes). Since some works have only the text components with no defining image features, a text-based spine recognition pipeline is combined with an image feature-based pipeline [24]. **Fig. 10 (b)** shows the detected books in different colors which are not of the same size, even though we got 100% accuracy **Fig. 10 (c)** shows the number of clusters into 14 pillboxes. **Fig. 10 (d)** shows the segmentation lines up to 18. The bright picture detects 100% and even segmentation one single line per pillbox, as shown in **Fig. 10 (e)**. **Fig. 10 (f)** shows the image's various shadows in the middle as bright, while the right and left sides are dark. The complete testing database contains over 1020 pictures of various medical industry medicine pill shelves and book stores used in the experiment to the introduced method. A collection of input data is shown in **Fig. 10**.

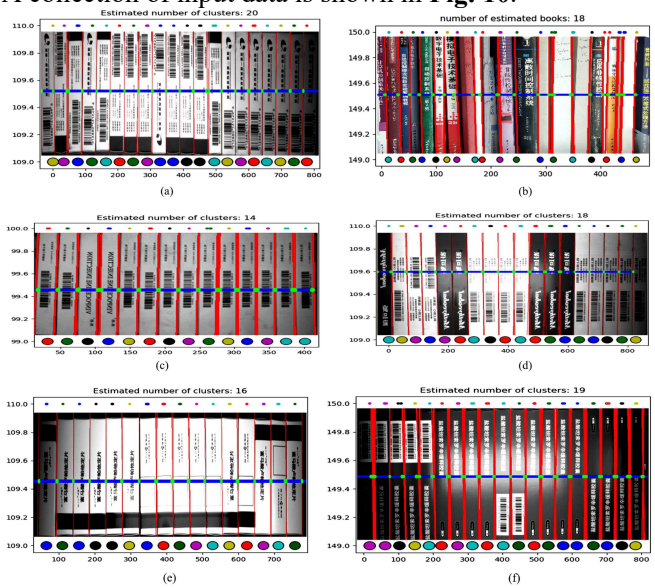

Fig. 10 The mean-shift algorithm is used to cluster the intersection points.

Following the segmentation operation, the intersection method is the approach for each image of a wide condition book's spine. The mean shift clustering technique then combines multiple intersect points to separate every object in such a row and uses an intersection points clustering method. As a result, we used the circle center clustering process to estimate the clustered points of various medical industry medicine pill shelves and book stores used in the experiment to the introduced method. A collection of input data is shown in **Fig. 10**.

Table. 2 Output of segmentation boundary images

No.	Segmentation images results			
	Samples	Results	Missing Result	Segmentation Images Rate (%)
1	110	110	0	100
2	95	95	0	100
3	135	134	1	99.5
4	85	84	1	99.5
5	150	105	0	100
6	125	125	0	100
7	120	119	1	99.5
8	95	95	0	100
9	105	105	0	100
Total	1020	1017	3	99.83

.

Table 2 shows an example of the suggested method's experimental data. On the testing data, the segmentation algorithm was performed with 95-100 % quality. Following the segmentation operation, the intersection method is the approach for each image of a wide condition book spine. The mean shift clustering technique then combines multiple intersection points to

separate every object in such a row and uses an intersection points clustering method. As a result, we used the circle center clustering process to estimate the clustered points. The segmentation approach performed 99.83 %. No. 1, the boundary segmentation images samples 110. No. 2, No. 5, No. 6, No. 8, No. 9, samples have segmented accurately. Here No. 3, one missed Because the hold picture was dark. No. 4, total image 85 also we obtained 84, the image is unclear. No. 7, is missing the detection line because the picture is much whiter. In our experiment, we successfully segmented over 1017 pillboxes and book spines out of 1020 images. Also, it results in a promising outcome at 99.83 percent, which outperforms the existing Hough line recognition with binary and color segmentation procedures. However, we have a few shortcomings in our drug box segmentation, which are hugely due to factors: first, the camera caption is poor, and second, the image cannot locate the segmentation axis because some pictures are dark and bright.

5. CONCLUSION

This project aims to reduce the arrangement time and solve the missing problem errors of medicine pillboxes or books on shelves using the object detection method. Initially, the input images have processed for correcting the angular orientations and the brightness quality. We use various types of spiral pillboxes and bookshelf images as input. The proposed method utilizes a vertical line detection formula to detect the segmentation boundary between the objects. Besides, the intersection points are created by calculating the midpoints of each segmentation line, which supports the correct counting of the object in the input image. Subsequently, the mean shift clustering algorithm accumulates multiple intersection points to separate each object in a row. We use a data set of 1020 images in various sizes and forms to conduct our experiment. In the image of our data set, some objectives on pill shelves had blank spaces between the object and few pixels are dark and bright, some angular condition pictures are also used in our experiment to get the best performance. The clustering intersection point and the counting outcome have 99.5% and 99.83% accuracy, respectively. The proposed methodology successfully recognizes pillboxes and books from a library bookshelf picture, according to simulation results. Future enhancement of our proposed method includes the detection of the horizontal segmentation boundary line.

REFERENCES

[1] Kim, M., Jeng, G.S., et al. Deep-learning im- age reconstruction for real-time photoacoustic system. IEEE transactions on medical imaging, vol. 39(11), pp.3379-3390, 2020.

[2] Nayan, Al-Akhir, Joyeta Saha, et al." Detection of Objects from Noisy Images." IEEE in 2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI), pp. 1-6. 2020.

[3] Benagi, Pratiksha and Kulkarni et al. Vector Based Object Identification in Spherical Images, Elsevier Procedia Computer Science, vol. 171, pp. 456–464, 2020.

[4] Jubair M I, Banik P. et al. A technique to detect books from library bookshelf image, IEEE 9th International Conference on Computational Cybernetics (ICCC). pp. 359-363, 2013.

[5] Quoc N H, Choi W H. A framework for recognition books on book-shelves, Springer International Conference on Intelligent Computing, pp. 386-395, 2009.

[6] Taira E, Uchida S, et al. Book boundary detection from bookshelf image based on model fitting, International Symposium on Information Science and Electrical Engineering, pp. 601-604, 2003.

[7] Prats M, Ramos-Garijo R. Autonomous localization and extraction of books in a library, 8th conference on intelligent autonomous systems, Amsterdam, The Netherlands, 2004.

[8] Yum H S, Hong M, Oh D I. Design and Implementation of a Book Counting System based on the Image Processing, Information Processing Society Software and Data Engineering Article, vol. 2, no. 3, pp.195-198, 2013.

[9] Duan X, Zhao Q. Identifying books in library using line segment detector and contour clustering, IEEE 7th International Conference on Computing and Convergence Technology (ICCCT), pp. 998-1003, 2012.

[10] Tabassum N, Chowdhury S. An approach to recognize book title from multi-cell bookshelf images, IEEE International Conference on Imaging, Vision Pattern Recognition (icIVPR), pp. 1-6, 2017.

[11] Kanburoglu A B, Tek F B. A Haar Classifier Based Call Number Detection and Counting Method for Library Books Ktphane Kitaplarinda Yer Numaralarini Bulmakve Saymaki in HaarSiniflandirici Tabanlibir Yntem.IEEE 3rd International Conference on Computer Science and Engineering (UBMK), pp. 504-508, 2018.

[12] Li L, Li D, et al. Extraction of road intersections from GPS traces based on the dominant orientations of roads. ISPRS International Journal of Geo-Information, vol. 6, no.12. pp. 403, 2017.

[13] Deutsch S, Medioni G G. Intersecting manifolds: detection, segmentation, and labeling, Twenty-Fourth International Joint Conference on Artificial Intelligence, pp. 3445-3452, 2015.

[14] https://docs.opencv.org/3.4/dd/dd7/tutorial morph lines detection.html. [Online]

[15] Sivkov, Stepan and Novikov. The algorithm development for operation of a computer vision system via the OpenCV library, Elsevier Procedia Computer Science, vol. 169, pp. 662 667, 2020.

[16] Zhao, Kai and Han, Qi and Zhang, et al. Deep hough transform for semantic line detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.

[17] Ni, Hui, Jinxia Xu, Ruoyue Wang, Jiangang Bi, et al." A Modified Mean Shift Method for Fault Region Extraction in Infrared Image." In Journal of Physics: Conference Series, vol. 1550, no. 2, p. 022010. IOP Publishing, 2020.

[18] Comaniciu, Dorin, and Peter Meer et al." Mean shift: A robust approach toward feature space analysis."IEEE Transactions on pattern analysis and machine intelligence, vol. 24, no. 5, pp. 603-619, 2002.

[19] Hongbin, Ma, and Ma Nan. Judgment motion by pressure clustering of limited nodes (three nodes) based on K-means, IEEE 12th International Conference on Electronic Measurement & Instruments (ICEMI), vol. 1, pp. 383-386, 2015.

[20] Moore A. K-means and Hierarchical Clustering, 2001. [Online] Available: http://www.cs.cmu.edu/ awm/tutorials

[21] Abualigah L, Gandomi A H, Elaziz M A, et al. Advances in Meta-Heuristic Optimization Algorithms in Big Data Text Clustering, Electronics, vol. 10, no. 2, pp.101, 2021.

[22] Nazari Z, Nazari M. et al. A Bottom-up Hierarchical Clustering Algorithm with Intersection Points. International journal of innovative computing, information and control, vol. 15, no. 1, pp. 291-304, 2019.

[23] Ohadi, N., Kamandi,SW-DBSCAN: A grid-based DB- SCAN algorithm for large datasets, IEEE 6th International Conference on Web Research (ICWR), pp. 139-145, 2020.

[24] Tsai S S, Chen D, et al. Combining image and text features: a hybrid approach to mobile book spine recognition, Proceedings of the 19th ACM international conference on Multimedia, pp. 1029-1032, 2011.