International Journal of Communication and Information Technology

E-ISSN: 2707-6628
P-ISSN: 2707-661X
www.computersciencejournals.com/ijcit
IJCIT 2023; 4(1): 18-25
Received: 13-10-2022
Accepted: 17-11-2022

Mayank Pandey
Software Engineer, Department of Computer Science Amity University, Uttar Pradesh, India

Anuwanshi Sharma
Ph.D., Scholar, Department of Basic and Applied Sciences, Galgotias University, Uttar Pradesh, India

Corresponding Author:
Anuwanshi Sharma
Ph.D., Scholar, Department of Basic and Applied Sciences, Galgotias University, Uttar Pradesh, India

Image segmentation techniques using python and deep learning

Mayank Pandey and Anuwanshi Sharma

DOI: https://doi.org/10.33545/2707661X.2023.v4.i1a.55

Abstract
Computer vision is one of the most important technology emerging in the IT market. Computer vision is a way to provide a powerful ability to see and interpret the real-time objects to the computer just like us or possibly even better than us. An image segmentation is the key processes in machine vision. It is the method of separate a high quality picture into non-overlapping multiple regions to get more precise and significant result. This technique provide us much more granular and reliable information about images like shape, color, and texture, etc. Image segmentation makes a huge impression in Image compression or Object detection techniques, Content-based image retrieval, Video surveillance, Medical imaging, and Traffic control system, etc. It is efficient to process small segments of images one by one to gather information and renounce the regions without information than to process an entire image for particular information. So, image segmentation is help us to dividing image into small segments for advance processing. We can divide a digital image using properties of images such as pixel values, intensity value, texture, shape, etc. So, we can achieve image segmentation by segmenting the images based on these properties. There are many techniques through which one can achieve image segmentation. Here, we discussed multiple techniques apply to achieve image segmentation and compare their benefits and drawbacks followed by the conclusion.

Keywords: Image segmentation, object detection, edge identification, thresholding, Clustering, deep learning

1. Introduction
An image is a visualized representation of something. It is a collection of various pixels which provide us much more useful results. It is a crucial and difficult technique in the field of digital image processing. Image segmentation is the process of extracting, analyzing, and comprehending relevant data from a single image or sequence of images by partitioning the image into non-overlapping multiple segments that contain considerably more useful information. We can divide the big images into small regions based on some digital image properties such as intensity values, color, called segments. It is inefficient to extract information from image by processing the whole image at once. So, image segmentation techniques much more suitable to reduce computing time [1].
Object detection and image segmentation are two different processes. We can group pixels with the same feature and build a box around each different object in an image. This process is known as object detection. This technique only provides us coordinate boundaries around a particular object belongs to various classes. But we need more deep information in an image to get more significant and accurate results. Let's consider an example, if we want to implement a process that enables the robot to catch a ball [3]. Let there are a ball and a bird in the view of the robot at the same time. So, we have to differentiate between them. Using segmentation techniques, we can form a mask for each object belongs to a particular class in the image. This process gives us much deeper data about images like shape, intensity value, color, etc.
Image segmentation is subjective in nature. The performance criteria will depend on what you want to observe in the image. Local segmentation and global segmentation are the two most common types. Local segmentation is concerned with a specific section of a picture for information, whereas global segmentation is concerned with splitting the entire image, which contains a vast number of pixels. There are two types of image segmentation based on how to classify objects in an image: Semantic segmentation, also called pixel-level classification and instance segmentation. Semantic segmentation is the task of gather segments together which having same object class whereas instance segmentation divides all objects even belongs to the same or different class.

Based on image attributes, image segmentation approaches can be divided into two sections:

1.1 A method based on the detection of discontinuities
This is a process of segmenting a picture into areas based on discontinuity. This is where edge detection based segmentation comes in, where edges generated by intensity discontinuity are recognized and linked to form region borders.

1.2 A technique based on affinity detection
This is a process of segmenting a picture into various divisions based on resemblance. Thresholding procedure, region expanding procedure, and region splitting and merging procedures are all included in this methodology. All of these divide the image into sections with similar pixel counts.

2. Related Work
In paper [5], the author discussed many different image segmentation methods like thresholding, region, and edge-based segmentation methods. Many edge detection, region, and threshold methods are tested on a benchmark and also discussed the performance concerning accuracy, error, and noise.

The paper [8] uses a survey research paper. They conclude that each technique apply to achieve image segmentation has its own uses and limitation. So, these segmentation methods can be performed based on the data we have and what type of result we want.

By merging deep learning and the SVM model, Feng Jiang *et al.* suggested a new method for detecting rice diseases in 2020. Because the region retrieved by this technique is big, the mean shift method was used as an image segmentation method. The time complexity of the model can be reduced by selecting the proper segmentation method. The collection consisted of 8,900 pictures in total. Other than typical approaches, three new elements of the crop leaf were evaluated, including shape features, area, roundness, and so on.

These shape features' eigenvalues were calculated and submitted to the SVM model. Three color spaces were used, including RGB, HSI, and YCbCr. Six separate levels were modeled in CNN. Three were convolutional, two were subsampling layers, and the final one was a feature layer that outputs a sigmoid activation function. Every layer generated feature maps with a specific pixel resolution, which were then passed on to the following layer. The SVM model was given the sigmoid activation function, three shape features, and a total of nine color spaces as input.

The weights between layers were initially chosen at random and then changed using a backpropagation technique. Because the penalty parameter (C) in SVM can remove certain redundant features, it was chosen. The accuracy results were compared using the grid search algorithm with different combinations of penalty parameter C and kernel functions. In comparison to the others, C = 1 and g = have the highest accuracy after performing 10-cross fold validation. The ROC curve was designed to fairly evaluate the model. Deep learning and SVM have a combined accuracy of 96.8%.

To accurately characterize radioactive wastes, the authors [12] devised a deep neural network (DNN) model utilizing ResNet. The model is based on six different categories of labeled data, four of which are radioactive wastes like vinyl, rubber, cotton, and paper, and two others like nothing in the image and no objects and no work tables. In all deep learning networks, the training and testing error rate increases as the number of layers grow. The Vanishing/Exploding gradient describes this behavior.

This occurs when the gradient associated with the deep learning network abruptly becomes 0 or excessively big. A dataset was built by capturing a video at a sorting worktable. The images were fine-tuned to a resolution that was adequate for training. The camera was afterward placed over the worktable for effective extraction.

3. Classification of image segmentation techniques
We can divide an image using properties of images such as pixel values, intensity value, texture, shape, etc. So, we can achieve image segmentation by segmenting the images based on these properties. There are many techniques through which one can achieve image segmentation with their own importance [2]. These approaches can be used to conduct segmentation on images. These techniques can alternatively be classified into three sections:

3.1 Structural Segmentation Techniques
These are picture segmentation process that rely on information about the structure of the image's necessary section.

3.2 Stochastic Segmentation Techniques
These are segmentation techniques that rely on the image's distinct pixel values rather than the region's structural information.

3.3 Hybrid Techniques
These techniques uses the concepts of both structural and stochastic segmentation process i.e. these uses discrete pixel and structural information together.

Now, we discussed and compared various process of image segmentation. The thresholding method, edge detection based procedure, region based procedure, clustering based procedure, watershed algorithm, interactive segmentation, color space based and regional convolution neural network segmentation are some of the popular image segmentation approaches. These strategies differ from one another in terms of the segmentation method they employ.

4. Techniques of image segmentation
4.1 Thresholding method
It is an effective and simple approach to image segmentation. This technique is mostly applied to grayscale images and uses grey level of pixels. We can totalize the pixels into a group having a value less than the threshold and the rest are grouped into another category. Using this technique, we can easily separate the object from its background if there is a large contrast between their pixel values. To determine the threshold value, we can use the peaks of the image histograms. The result of thresholding is an image that has exactly two colors, usually black and white, called a binary image [4].

Threshoulding are classified into three categories: Global Thresholding (describing a single appropriate threshold value, T to classify exactly two regions), Multiple Thresholding (describe more than one value of threshold to classify more than two regions), and Variable thresholding (threshold value can vary over the image).

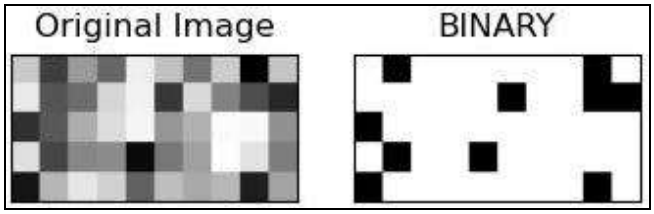

Fig 1: Binary Thersholding

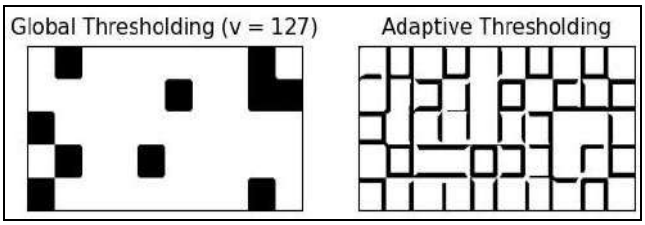

Fig 2: Adaptive Thersholding

As a threshold value, we used a global value. It might not be ideal in some situations, such as when the image has varying flash conditions in different locations. As a result, we can employ adaptive thresholding. The method calculates the threshold for a small area of the image in this way.

4.2 Region based segmentation

Region-based segmentation is an efficient procedure which categories an image into sub-regions with same characters. Region growing and region merging and splitting are two basic methods to proceed towards region-based segmentation.

4.2.1 Region growing method

In this procedure to segmentation, we inspect adjacent pixels of the initial seed point. The seed point is an initial position in an image. This method is also known as pixel-based image segmentation since it consists of a selection of seed points. In the region growing method, a sub-region is formed with a set of pixels near to seed point having similar characteristics to the seed point. The sub-region append from seed points to neighboring points depending upon some special criteria such as pixel value, intensity, texture, or color. This method is better than the edge detection techniques for images with high values of noise [5].

Suppose we have an image given below

Table 1: Demonstrating Region-based Segmentation

10	25	28	19	22	43	23
12	22	17	22	12	10	17
14	18	20	28	23	18	12
19	28	34	(24)	10	12	34
31	12	23	20	11	43	48
28	15	12	19	43	54	47
25	13	11	17	34	48	34

The seed of this image is 24 with oval shape. Now, apply region growing method with membership criteria difference between two pixel value less than or more than 5 to the seed point. The result image from region growing method given below.

Table 2: Demonstrating Region-based Segmentation

10	25	28	14	22	43	23
12	22	15	24	60	70	17
14	18	40	28	23	18	11
59	48	44	(24)	30	29	34
31	12	43	20	11	43	48
28	15	12	19	43	54	47
25	13	11	17	34	48	34

4.2.2 Region splitting and merging methods

This approach to segmentation based on two basic techniques: splitting and merging. It is an alternative to the region growing where we start the analysis from the seed point. Here, we start with the entire image and divide the regions that do not fulfill the membership criteria. Region merging has just opposite function. The analysis starts with small regions (such as 2*2 or 4*4 sub-region) and merges the regions that have homogeneity.

4.3 Interactive image segmentation (watershed algorithm)

Self-regulating segmentation is a complex problem. And even after modeling automatic segmentation, some human involvement is required to correct the abnormality in the segmentation.

The interactive segmentation approach based on the proper user's participation in the segmentation process. The user provides some indication for segmenting the image which directs to adequate results. The semiautomatic or interactive segmentation method uses human intelligence as an auxiliary input and offers the segmentation problem more flexible.

An image with many shades of gray can be observed as a topographic surface where high light indicates hills while low light expresses valleys. You start loading every separate valley with various labels. As the water grows, water from different valleys with several colors will begin to mix. To evade that, you build obstacles in the spots where water blends. You resume the task until all the peaks are under the water. Then the boundaries you sketched provide you the segmentation outcome. This is the philosophy behind the watershed. There are also various diverse algorithms to measure the watersheds. The user can implement distinct procedures to use the watershed system for image segmentation [6].

However, due to noise or other defects in the image, this approach results in an over segmented result. You utilize a marker-based watershed algorithm to determine which valley points should be consolidated and which should not. What we do is to provide various tickets for our object we acknowledge. Label the domain which is the foreground, label the quarter which is the background and label the region which we are not assured of anything with 0. This is the marker after that implement the watershed algorithm. Then our marker will be refreshed with the tickets we provided, and the edges of items will have a grade of -1. Marker-based watershed makes use of definite marker points which have been either explicitly specified by the user or managed automatically.

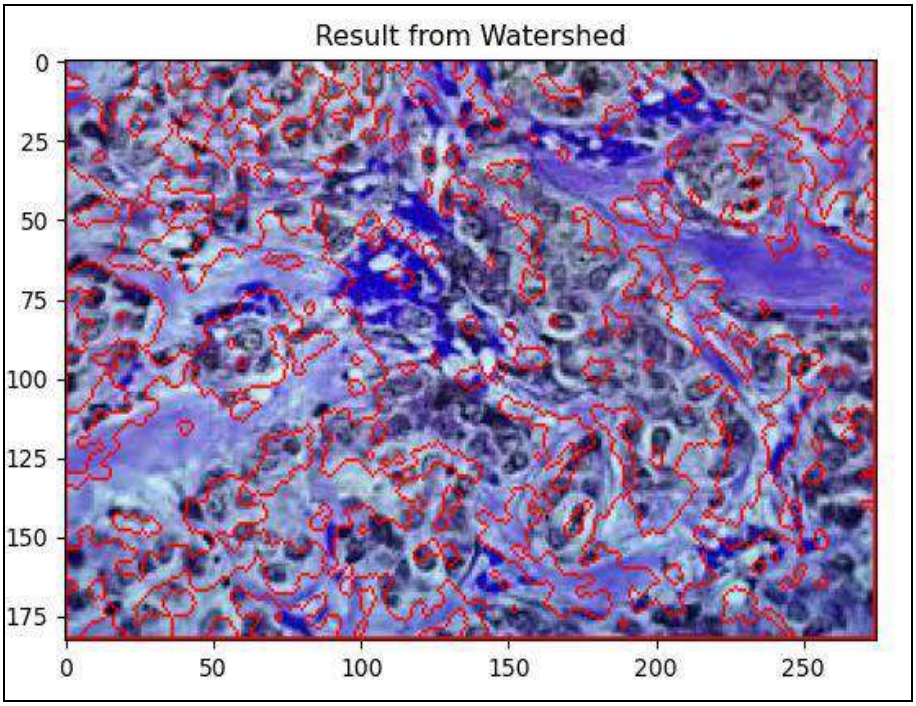

Fig 3: Input image for watershed algorithm

4.4 Edge detection

This procedure are efficient and well-established methods of image segmentation. This method attempts to resolve the image segmentation by recognizing the edges within various regions that have an active transformation in intensity. The outcome of the method is a binary picture.

A single intensity value does not provide sufficient data about edges. The edge-based segmentation algorithms rely on active switching of intensity values in a picture. In this procedure, all of the edges are first recognized, then linked together to generate the object edges, which are then used to segment the desired regions. Two basic edge-based segmentation approaches are grey histograms and gradient-based algorithms. The primary edge detection techniques such as the sobel operator, canny operator, and pewit, among others, can be utilized to identify the edges unit. These are the fundamental techniques based on discontinuity detection.

Prewitt proposed the pewit edge detection method in 1970. It is applied to an image to detect its horizontal or vertical boundaries. Using the contrast between the identical pixel intensity of an image, edges can be detected. Imitative masks are utilized for boundary detection. These masks have attributes like the entirety of the mask should be equal to zero, should have an inverse sign and large weight so we can efficiently identify the edges.

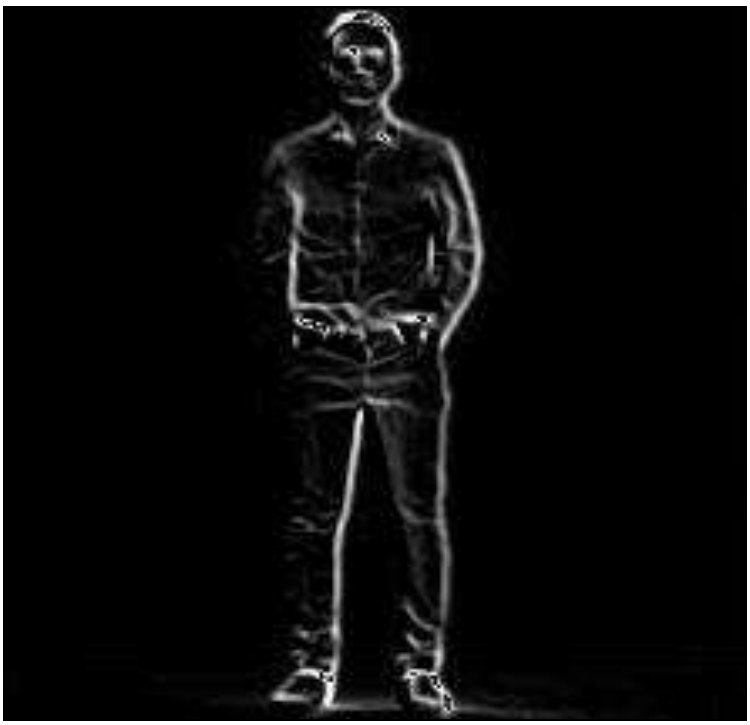

Fig 4: Prewitt segmentation

The Canny edge detector is an operator that works on a multi-stage algorithm to identify an extensive chain of edges in images. It was proposed by John F. Canny in 1986 [7].
The Canny edge detection method has the following different steps:

1. Implementing the Gaussian filter to eliminate the noise.
2. Determining concentration gradients
3. Use non-maximum suppression to clear artificial acknowledgment to edge detection
4. Identify potential edges and mark bounds by hysteresis

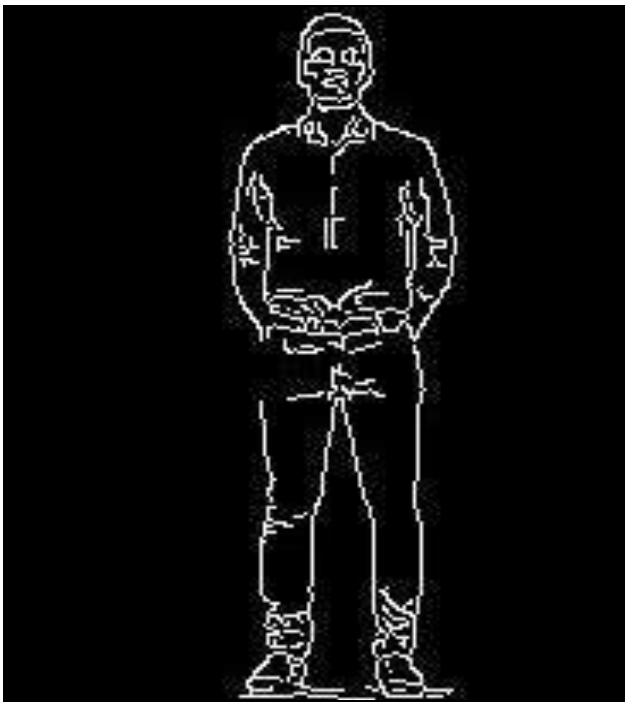

Fig 5: Canny Edge Detector

The Sobel edge detection method is pretty alike to the prewitt method. The principal variation between these two methods is the coefficient of masks isn't fixed. We can modify it according to our necessity unless they don't disrupt any feature of derivative marks.

Fig 6: Sobel Edge Detection

E. Color space based image segmentation

Color space has an active position in image segmentation techniques. RGB and HSV are different types of color space. RGB colors are described in expressions of their red, green, and blue elements in standard color space and as a tuple of three elements. Every element has a value between 0 and 255, where (0, 0, 0) denotes black and (255, 255, 255) expresses white.
HSV is the information of hue, saturation, and luster, which are especially valuable for classifying inconsistency in images. But RGB color space is inaccurate in segmentation techniques. It is irrelevant for human hand tracking and gesturing. A heterogeneous technique (RGB-HSV) is introduced that can provide comprehensive data still in a ruined background and can provide enhanced color conserve information which is extremely beneficial for hand motioning.

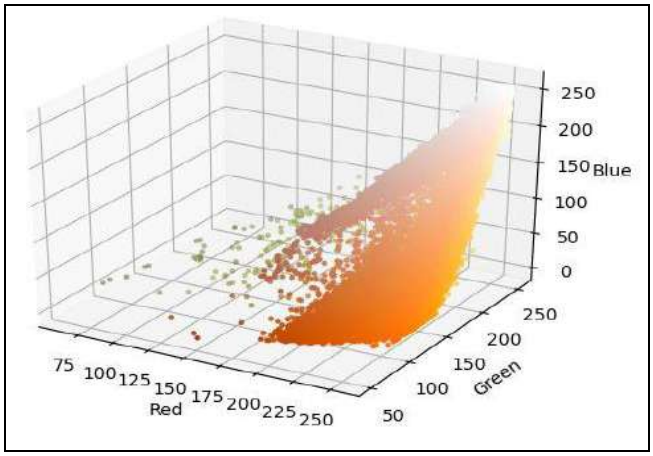

Fig 7: RBG Plot

From the above figure, we can understand that the orange part of the image spread over nearly the whole range of red, green, and blue values. Segmenting fruit using spectra of RGB values is not be an easy task because of parts of the fruit extent over the almost entire plot.

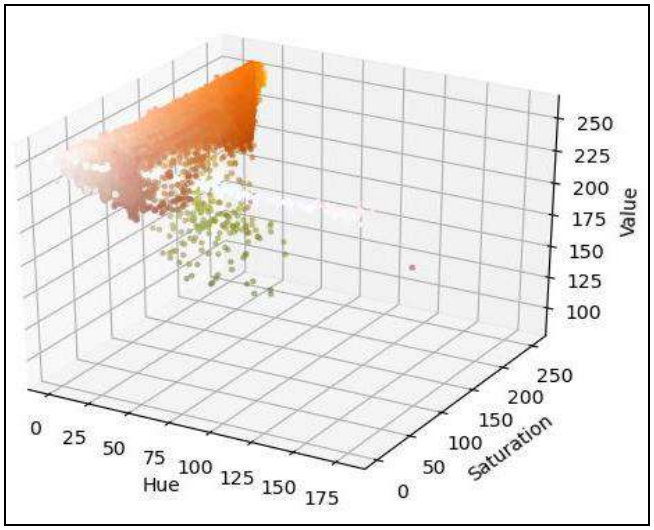

Fig 8: HSV Plot

Fruit's oranges are more confined and visually divisible in HSV space. The saturation and value of the fruit are differing, but they are established on a small scale with the hue axis. This is the primary target for segmentation.

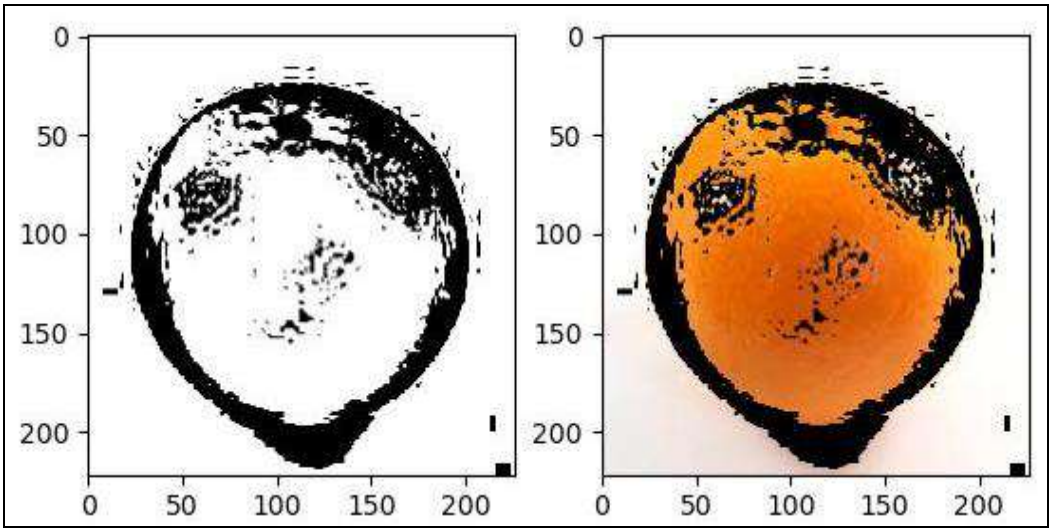

Fig 9: Output image for color space-based segmentation

F. Clustering based image segmentation

The clustering-based methods partition the image into clusters occupying pixels with alike properties. The hierarchical approach and the partition approach, two primary methods for clustering that uses the theory of trees and optimization approaches iteratively to reduce an objective function respectively. There are many techniques for clustering such as Kmeans clustering, mean-shift clustering, etc. [8].

Kmeans clustering is a method to analyze or to gather the objects in an image based on characteristics into K groups. Here, K represents a positive number. By reducing the interval between data and the analogous cluster centroid, classification is done. The primary goal of clustering to classify given data in such a way that related objects fall in one cluster and objects with divergent properties belongs to a different cluster. The K-means clustering algorithm has the following three steps:

1. Initialize the position of K or centroids.
2. Data is allocated to its most neighboring centroid
3. The location of the centroid is refreshed by the averages of the data points allocated to that cluster.

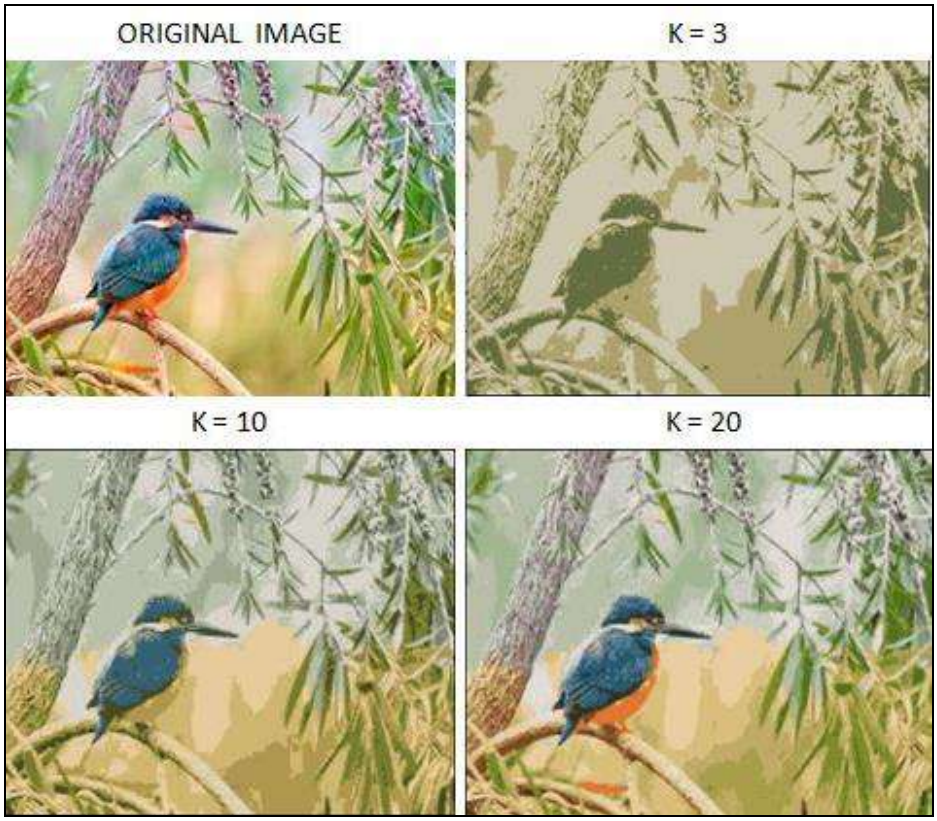

Fig 10: (a) Original Image, (b) Clustering Output when k = 3, (c) Clustering Output when k = 10, (d) Clustering Output when k = 20

The above image showing the result of K means clustering when applied to the original image. Here, you can easily see that as we increase value of K then image is more and more sharp and clear but the computing time will increase significantly.

Now, to understand K means more effectively we remove the 2nd k cluster from the above image.

Fig 11: Output Image When 2nd Cluster Removed

In the above image, the black colored portion shows segment of image where 2nd cluster was located earlier. In other word, black portion has similar pixels and other features.

4.5 Mask R-CNN

Mask R-CNN is proposed by researchers of the Facebook AI Research (FAIR) that generate pixel-wise mask to all articles in an image. The main goal to develop this architecture is to solve fast and accurate image segmentation problem in computer vision. YOLO (You Look Only Once) is one of the best and fastest methods to detect objects, based on R-CNN architecture.

Mask R-CNN is a branch of the popular CNN architecture and uses to detect the objects in an image. Mask R-CNN method produces three elements for each object that are its class, box around the object, and object mask (shape) as well. Mask R-CNN has the following step:

1. Passing an image to the ConvNet, which produce the characteristic map for that image
2. Region proposal network (RPN) returns the object patterns with their object score depending upon number of objects.
3. An RoI pooling layer is used on these patterns to reduce all the pattern to the identical size
4. Finally, the connected layer is used to classify these patterns and produce output with a box around the object and the mask [9].

Till now, we discussed many methods to achieve image segmentation. But these are only some pages from a book. There are many more segmentation techniques like partial differential equation-based segmentation method (relevant for critical time applications and used to diminish the noise from the image and to identify the edges), Texture based segmentation (uses texture analysis), etc. Now, we compare these techniques trough a table given below.

Table 3: Applications and limitations

Name of image segmentation technique	Applications	Limitations
Threshoulding method	• Simple • Used when pixels of object and background have enough contrast	• Difficult to segment when there is no sharp difference in their pixels
Region based segmentation	• Accurate and fast • Simple to implement	• Computationally expensive • Sensible to noise • No global view
Watershed image segmentation	• Favorably used on a variety of data for segmentation • Researchers are obtained the brain from MRI data	• Return over-segmented images without proper use of parameters
Edge detection	• Suitable for images with contrast between	• Difficult to compute if image has too many edges
Color space based image segmentation	• color spaces perform better with proper parameters	• Non-removable singularities • Difference between color is not linear
Clustering based image segmentation	• Excellent and suitable approach with many advantages	• Time consumption is high • Kmeans not good for non-convex clusters
Mask R-CNN	• Simple, flexible and accurate • Currently popular	• Too much training time

5. Advantages of image segmentation

There are many fields which enhance their application with the help of image segmentation such as medical, traffic, and monitoring, etc. Now, we discussed some major application of image segmentation which are as follows:

5.1 Remote sensing

It is used to identify energy which mirrors back by earth and retrieve information from it. It is a developing field of recognizing the data. The remote sensing is capable of observing the shifts in the environment like deforestations, ecosystem degeneration, and nitrogen and carbon recycling, etc.

5.2 Medical imaging

Medical imaging is also a growing field that helps us to generate the visual representation of human internal organs and to examine the internal parts of the human body such as bones etc. We can also diagnose any hidden disease more accurately using this technique.

5.3 Aerial imaging

We are mostly using drones for monitoring areas with a high crime rate and sometimes at wartime or riots. With the help of this technology, we easily get information of any place from anywhere. Image segmentation helps a lot in this field to detect a specific human using face detection or any object.

5.4 Smart traffic control

Today, this is not an implementing task. But many types of research are going on this hot topic. This helps us to control traffic according to certain conditions without human intervention. With image segmentation and digital image

processing, we can generate a highly efficient and intelligent traffic controller.

6. Conclusion and future scope

We reviewed and compared various image segmentation techniques in detail, in this study of image segmentation techniques. These techniques have their own uses and limitations. Region-based and edge detection is very helpful in analyzing satellite images and segmenting roads and other objects. Image segmentation also has a valuable role in artificial engineering. K means clustering and color based clustering can be used to detect cancerous cells in cell clusters. In this paper, we noticed that each method has been used based on data and its quality provided. Different techniques are suitable for various tasks. But from the above analysis, no single method is satisfactory for every image type. Some are suitable for black and white images and some for colored. It is a popular issue in science right now, with applications ranging from computer vision to medical imaging to traffic and video monitoring. It faces a difficult future due to the growing need for picture segmentation in numerous applications.

7. Ethics approval and consent to participate

Not applicable.

8. Conflicts of interests

The authors declare that they have no conflicts of interests.

9. Acknowledgments

I would like to convey my heartfelt gratitude to everyone specially my parents who has aided me greatly and supported me at all times, as well as provided me with vital inspiration and advice in my search for knowledge and completion of this work. I give them a debt of gratitude.

10. References

1. Maheswari S, Korah R. Review on image segmentation based on color space and its hybrid. International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Kumaracoil; c2016. p. 639-641.
2. Song Y, Yan H. Image Segmentation Techniques Overview. Asia Modelling Symposium (AMS), Kota Kinabalu; c2017. p. 103-107.
3. Varshney SS, Rajpal N, Purwar R. Comparative study of image segmentation techniques and object matching using segmentation. International Conference on Methods and Models in Computer Science (ICM2CS), Delhi; c2009. p. 1-6.
4. Sahu S, Sarma H, Jyoti D. Bora. Image Segmentation and its Different Techniques: An In-Depth Analysis. International Conference on Research in Intelligent and Computing in Engineering (RICE), San Salvador; c2018. p. 1-7.
5. Zou Y, Liu B. Survey on clustering-based image segmentation techniques. IEEE International Conference on Computer Supported Cooperative Work in Design (CSCWD), Nanchang; c2016. p. 106-110.
6. Sevak JS, Kapadia AD, Chavda JB, Shah A, Rahevar M. Survey on semantic image segmentation techniques. International Conference on Intelligent Sustainable Systems (ICISS), Palladam; c2017. p. 306-313.
7. Chauhan AS, Silakari S, Dixit M. Image Segmentation Methods: A Survey Approach. International Conference on communication systems and network technologies, Bhopal; c2014. p. 929-933.
8. Taneja A, Ranjan P, Ujjlayan A. A performance study of image segmentation techniques. International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), Noida; c2015. p. 1-6.
9. Sivakumar V, Murugesh V. A brief study of image segmentation using Thresholding Technique on a Noisy Image. International conference on information communication and embedded systems (ICICES2014), Chennai; c2014. p. 1-6.
10. Chandra JN, Supraja BS, Bhavana V. A Survey on Advanced segmentation techniques in image processing applications. IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Coimbatore; c2017. p. 1-5.
11. Jrad MS, Oueslati AE, Lachiri Z. Image segmentation based thresholding technique: Application to DNA sequence scalograms. International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Monastir; c2016. p. 319-324.
12. Jeong-Guk Kim, *et al*. A study on object recognition using deep learning for optimizing categorization of radioactive waste. Progress in Nuclear Energy; c2020. p. 130.