ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jun1813

A Survey of Image Segmentation Using K-means Clustering

Koushik Mukherjee[1]; Soumen Bhowmik[2]

[1]Research Scholar M. Tech, CSE, Bengal Institute of Technology and Management Santiniketan, India
[2]Head of the Department, CSE, Bengal Institute of Technology and Management Santiniketan, India

Publication Date: 2025/07/21

Abstract: In this survey paper we have tried to identify the necessity and advantages and disadvantages of different approaches of different researchers. We know the process of grouping objects or datasets with comparable data points is called clustering. Groups of items are as similar as they can be. Each group's points are as dissimilar as possible. One well-liked algorithm for classifying and dividing data is K-means clustering. It facilitates the separation of objects from their surroundings. According to K-means clustering, every feature point has a distinct location in space. In a multidimensional measurement space, a predetermined number of cluster centers are chosen at random by the fundamental K-means algorithm. The cluster with the closest mean vector is assigned to each pixel in the picture. This procedure continues until the cluster mean vector positions between iterations vary as little as possible. The K-means algorithm is significantly impacted by the initial starting points. Initial clusters are randomly generated by K-means. When these starting points are chosen close to the final solution, K-means can usually find the cluster center effectively. Otherwise, it could lead to poor clustering results. K-means clustering is one method for classifying data. The K-means function yields the cluster index for each observation following the division of the data into k distinct clusters.

Keywords: *K-means Clustering; Segmentation; Pixel.*

How to Cite: Koushik Mukherjee; Soumen Bhowmik (2025) A Survey of Image Segmentation Using K-means Clustering. *International Journal of Innovative Science and Research Technology*, 10(6), 3293-3295.
https://doi.org/10.38124/ijisrt/25jun1813

I. INTRODUCTION

One popular technique for locating the pixels in an image for decision-making is image segmentation. It separates an image into multiple discrete regions, maintaining a high contrast between them while guaranteeing that the pixels in each region are extremely similar. This method is helpful in a number of domains, including pattern recognition, traffic signals, image processing, and healthcare. Image segmentation can be done in a number of ways, such as using neural networks, thresholds, edges, and clustering. Among these strategies, the clustering approach is among the most successful. K-means clustering, mountain clustering, fuzzy C-means clustering, and subtractive clustering are among the various kinds of clustering techniques. Clustering outcomes can vary depending on the ideal number of clusters. [1]

These days, image segmentation is a crucial medical method for distinguishing a particular region of interest from its surroundings. Medical images are segmented using various methods, and the outcomes of this process are utilized for additional medical research. In their original format, medical images are displayed as numerical arrays. The values of significant physical quantities that illustrate the variations among different body parts are represented by the numbers displayed on the computer. [3]

Processing and analysis of medical images combine data from various imaging techniques, transform raw images into a measurable format, and extract important qualitative information to aid in diagnosis.[2] Finding the edges of objects, like organs or areas of interest, within images is the main goal of image segmentation, a crucial component of medical analysis. Shape analysis, volume change detection, and the creation of an exact radiation therapy treatment plan are all made possible by the segmentation results. [5]

II. SUGGESTED K-MEANS CLUSTRING METHOD

Training data from various classes, with k denoting the number of clusters, make up the inputs. Determine the centroids independently for every class. Determine the centroid for a set of points x1, x2,..., xn by calculating the Euclidean distance between each point xi and the closest centroid. The Euclidean distance It is possible to calculate the distance (d) between two points in Euclidean space, p = (p1, p2,..., pn) and q = (q1, q2,..., qn). The cluster whose mean has the smallest squared Euclidean distance should be given each point.” [4].

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jun1813

$$d(\mathbf{p},\mathbf{q}) = d(\mathbf{q},\mathbf{p}) = \sqrt{(q_1-p_1)^2 + (q_2-p_2)^2 + \cdots + (q_n-p_n)^2}$$
$$= \sqrt{\sum_{i=1}^{n}(q_i-p_i)^2}.$$

To indicate which mean is closest, assign each point to the cluster whose mean has the smallest squared Euclidean distance. Until the clusters are formed, repeat the same procedures for every cycle. The process of splitting a digital image into several segments—groups of pixels known as super pixels—is known as image segmentation in computer vision. Simplifying an image's representation or transforming it into a more meaningful and comprehensible format are the goals of segmentation. To identify objects and their borders, such as lines and curves, in images, image segmentation is frequently utilized.

The process of giving each pixel in an image a label and making sure that pixels with the same label have similar characteristics is known as image segmentation. A collection of segments that collectively depict the full image or a number of outlines extracted from it is the result of image segmentation. Common or measurable attributes, like color, brightness, or texture, are shared by pixels in a given area. The same characteristic varies significantly in the surrounding areas. Clustering is the process of grouping samples so that they are comparable to one another within each group. We refer to these groups as clusters. In data mining, clustering is a technique used for tasks like image analysis, pattern recognition, statistical analysis, and other related fields. Although there will be less similarity between groups, a good clustering technique will create groups that are highly similar to one another. The method's implementation and the similarity metric it selects determine how well the clustering results turn out. By evaluating a clustering method's capacity to identify some or all hidden patterns, it evaluates how effective it is. One of the most important steps in the analysis and comprehension of images is image segmentation. It is considered one of the oldest and most challenging problems in image processing and is crucial to the field. The process of grouping and differentiating objects based on shared characteristics is called clustering. Using clustering techniques, pixels with similar characteristics are grouped together to form various clusters depending on the similar [7].

III. RESULT FROM EXPRIMENT

Following results are obtained by the different authors [8] [9] [10] after segmenting image data using an iterative approach. The random selection of initial centroids at the beginning of the algorithm can result in different clusters, as the k-means algorithm might become trapped in a local optimum and not reach the global optimum. In the Figure 1 and Figure 2, the authors capture an aerial image and perform image segmentation using K=2 and K=4 .

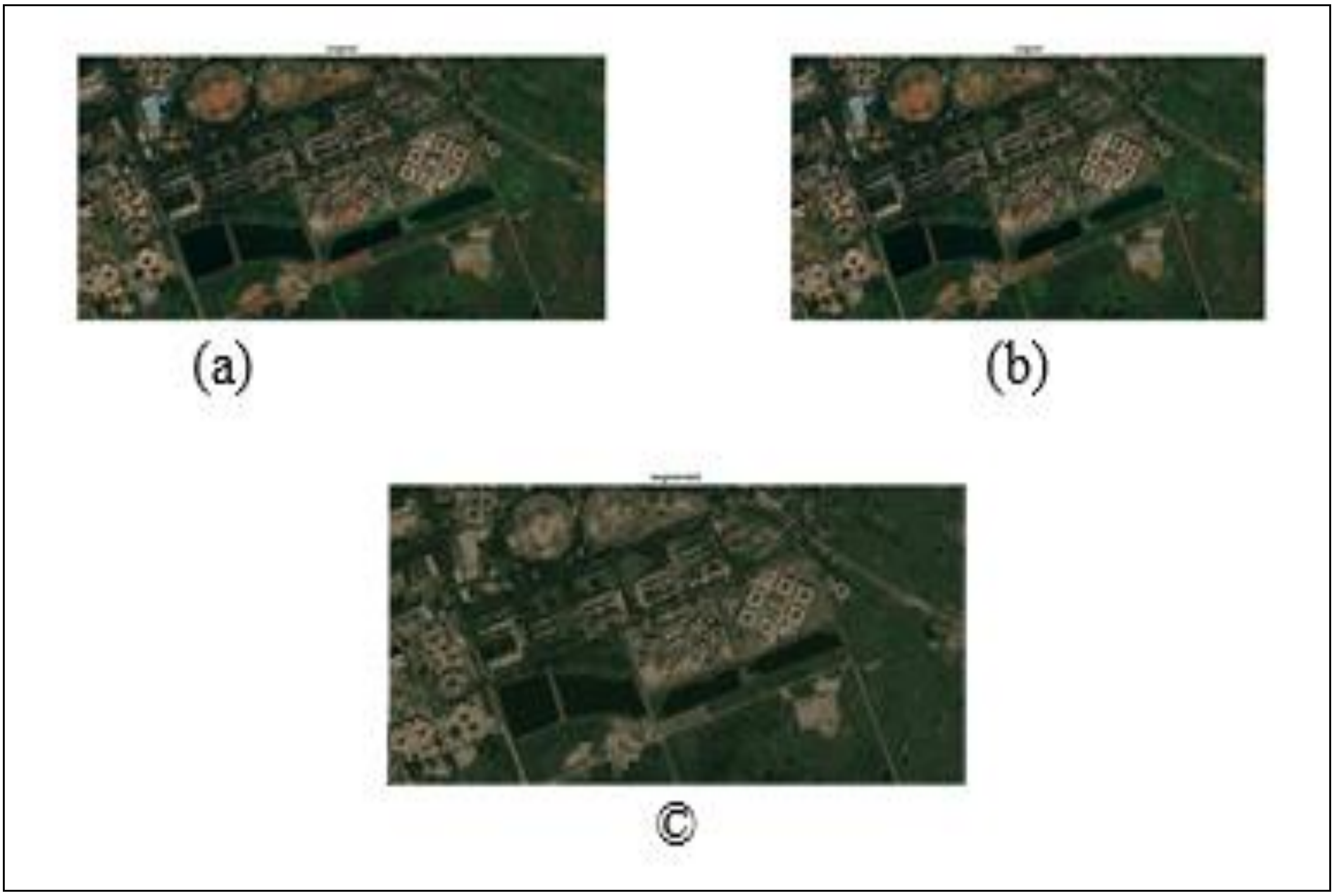

Fig 1: (a) Original Image. (b) Segmented Image with K=2. (c) Segmented Image with K=4

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jun1813

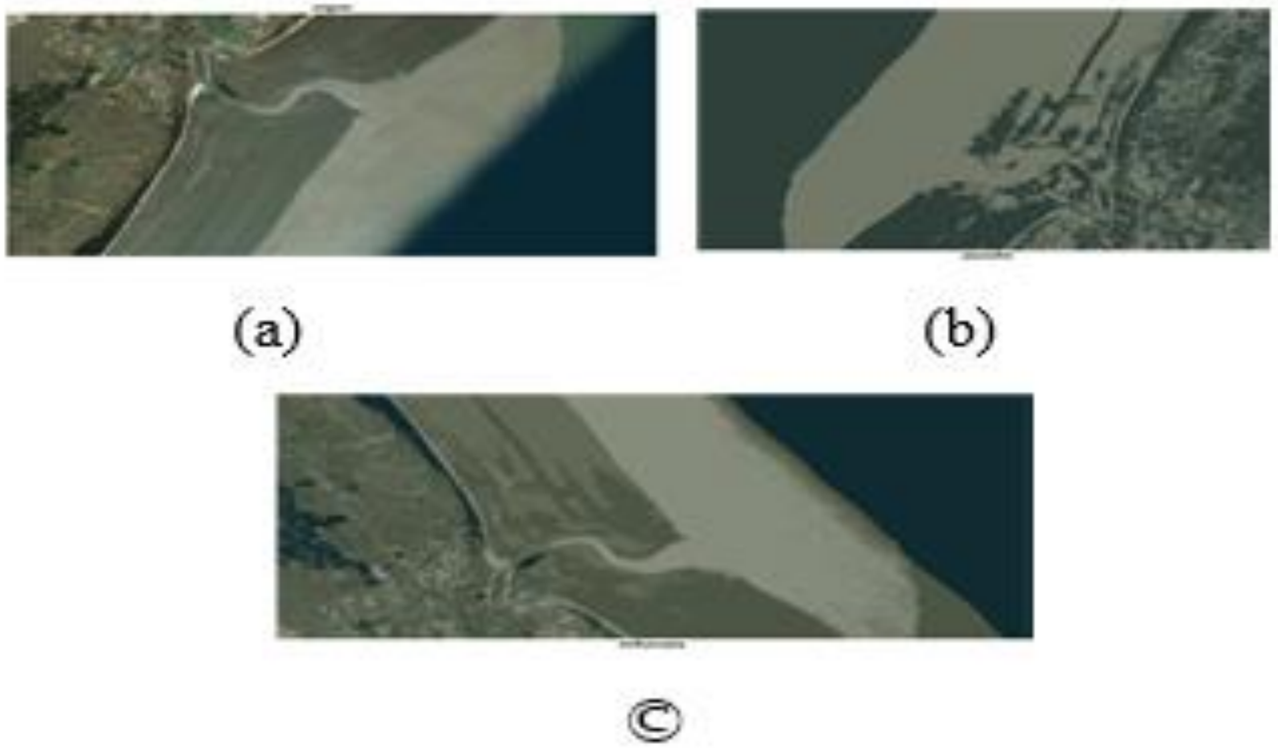

Fig 2: (a) Original Image. (b) Segmented Image with K=2. (c) Segmented Image with K=4.

The K-MEANS algorithm aims to minimize the number of distance calculations and time required, while also enhancing accuracy [10]. An increase in the number of pixels in an image leads to a longer execution time. Two Arial images, one have a higher pixel count and another having lower pixel cont are segmented using the K-means clustering algorithm with varying K values.

IV. CONCLUSION

Segmentation is employed for quantifying classes in aerial images, allowing us to categorize features such as buildings, soil, and water. In this context, K signifies the clusters and categories associated with each type of area, allowing us to determine the percentage of a specific area represented in that aerial image. One drawback of using the K-means algorithm for image segmentation is present. The number of calculations for distance functions is higher. With this observation we can conclude that much more research is required to fulfill all the aspect of the quality of the segmentation.

REFERENCES

[1]. A.T. Michailovich, Oleg, Yogesh Rathi, “"Image segmentation using active contours driven by the Bhattacharyya gradient flow.,” IEEE Trans. Image Process., vol. 16.11, pp. 2787–2801, 2007.
[2]. F.jian. and M.R.Frater, “Threshold-based image segmentation through an improved particle swarm optimization,” in Digital image computing techniques and applications (DICTA), 2012.
[3]. P. Y. and G. R. Reddy, “Segmentation of noisy images using improved distance regularized level set evolution,” in IEEE International conference on circuits, power and computing technologies, 2017.
[4]. A. S. Jumb, Vijay, Mandar Sohani, “"Color image segmentation using K-means clustering and Otsu’s adaptive thresholding.",” Int. J. Innov. Technol. Explor. Eng., vol. 3.9, pp. 72–76, 2014.
[5]. E. Deligiannidis, Leonidas, and Hamid R. Arabnia, “Emerging trends in image processing, computer vision and pattern recognition.,” in Morgan Kaufmann, 2014.
[6]. S. B.-A. Al Bashish, Dheeb, Malik Braik, “Detection and classification of leaf diseases using K-means-based segmentation and.,” Inf. Technol. J., vol. 10.2, pp. 267–275, 2011.
[7]. H. K. H. Wong, Stephen TC, “Design methods and architectural issues of integrated medical image data base systems.,” in Computerized Medical Imaging and Graphics, 1996, p. 285–299.
[8]. K. V. and B. K. Singh, “An enhancement in adaptive median filter for edge preservation,” in ScienceDirect, Computer science procedia, 2015, pp. 29–36.
[9]. P. Vijayalakshmi, “Segmentation of Brain MRI using K-means Clustering Algorithm,” Int. J. Eng. Trends Technol., pp. 113–115, 2011.
[10]. K. M. and Y. J. chanu N. Dhanachandra, “ Image segmentation using K-means clustering algorithm and subtractive clustering algorithm”, 2015, pp. 764-771.,” in ScienceDirect, Computer science procedia, 2015, pp. 764–771.