# Optimization techniques applied on image segmentation process by prediction of data using data mining techniques

**Ramaraj Muniappan[1], Srividhya Selvaraj[2], Rani Vanathi Gurusamy[3], Velumani Thiyagarajan[1], Dhendapani Sabareeswaran[4], David Prasanth[5], Varadharaj Krithika[6], Bhaarathi Ilango[1], Dhinakaran Subramanian[1]**

[1]Department of Computer Science, Rathinam College of Arts and Science, Coimbatore, India
[2]Department of Computer Science, KPR College of Arts Science and Research, Coimbatore, India
[3]Department of Computer Science, Sri Ramakrishna CAS for Women, Coimbatore, India
[4]Department of Computer Science, Govt. College of Arts and Science for Women, Tiruppur, India
[5]Department of Computer Science, Shri Nehru Maha Vidyalaya College of Arts and Science, Coimbatore, India
[6]Department of Computer Science, Akshaya College of Engineering and Technology, Coimbatore, India

**Article Info**

***Article history:***

Received Aug 1, 2024
Revised Oct 18, 2024
Accepted Nov 20, 2024

***Keywords:***

Classification
Clustering
Color image segmentation
Density-based spatial clustering of applications with noise
Fuzz logic
Pattern mining
Performance metrics

**ABSTRACT**

The research work presents an enhanced method that combines rule-based color image segmentation with fuzzy density-based spatial clustering of applications with noise (FDBSCAN). This technique enhances super-pixel robustness and improves overall image quality, offering a more effective solution for image segmentation. The study is specifically applied to the challenging and novel task of predicting the age of tigers from camera trap images, a critical issue in the emerging field of wildlife research. The task is fraught with challenges, particularly due to variations in image scale and thickness. Proposed methods demonstrate that significant improvements over existing techniques through the broader set of parameters of min and max to achieve superior segmentation results. The proposed approach optimizes segmentation by integrating fuzzy clustering with rule-based techniques, leading to improved accuracy and efficiency in processing color images. This innovation could greatly benefit further research and applications in real-world scenarios. Additionally, the scale and thickness variations of the present barracuda panorama knowledge base offer many advantages over other enhancement strategies that have been proposed for the use of these techniques. The experiments show that the proposed algorithm can utilize a wider range of parameters to achieve better segmentation results.



***Corresponding Author:***

Ramaraj Muniappan
Department of Computer Science, Rathinam College of Arts and Science
Coimbatore-21, Tamil Nadu, India
Email: ramaraj.phdcs@gmail.com

## 1. INTRODUCTION

The unsupervised learning methods employed can challenging to the cluster data points based on particular characteristics. Clustering algorithms, including K-Means and hierarchical, are extensively utilized, offering effective solutions to these challenges through the application of clustering concepts [1]. They cannot bunch together in light of their different densities. The proposed clustering algorithm will be focused on the hypothesis that high-density locations are clustered around one another to constitute a single-color group when isolated by thickening districts [2]. This technique analyzes the local density of data points within clusters, enabling the identification of clusters in large spatial datasets based on color data. It helps pinpoint the cluster group most likely to form [3]. The two key parameters required are Epsilon and MinPoints, which define the radius and density threshold for evaluating the neighborhood of each data point. These parameters help in accurately assessing the cluster's density by determining how many neighboring points fall within a specified distance around each point.

MinPoints is the base number of points that need to be in a circle for the highlighted information to be considered the center point in the list [4]. In density-based spatial clustering of applications with noise (DBSCAN), the minimum number of data points required within a hypersphere is termed MinPoints. The proposed algorithm categorizes data points within an epsilon radius as core, border, or noise points. Core points are those surrounded by at least MinPoints within their neighborhood radius [5]. A point is viewed as the boundary point in the event that the quantity of points of interest is not an exact quantity of points, and a point is viewed as commotion on the off chance that there could be no different data of interest inside an epsilon span than there is in any data of interest [6]. Core points, which are depicted by the color red, are all data points that have at least three points in the circle, including themselves. The Euclidean distance is used by DBSCAN to locate data points in space. But other distances are calculated by the colors in an image, such as the wide circular distance, which can used to locating geographical data [7]. Even though, it has need to do it multiple times in our calculations, it only needs to look at the entire dataset once [8].

## 2. REVIEW OF LITERATURE

Numerous approaches have been investigated for image segmentation, with much of the research centered on different segmentation technologies. The innovative and unconventional work currently being conducted in image processing is particularly intriguing. Yang *et al.* [8] developed a novel and efficient K-hyper line in clustering based color image segmentations (CBCIS) strategy that relics resistant to vicissitudes in illumination, it has been demonstrated for color image classification as a hyper line clustering solution for color image classification in the visual field, which is used to identify color images in the visual field. Ramaraj and Niraimathi [9] presented a few issues within the variety image division. With the five issues that were discussed in the color image segmentation, proposed solutions are taken into consideration. Zhu *et al.* [10] summed up the various approaches to image segmentation that have been used in this situation. Yan *et al.* [11] described a new framework for comparing various clustering methods for segmenting pixels of an images.

Chen *et al.* [12] has proposed new adaptive methods for estimating initial parameters in image segmentation by utilizing an improved k-means clustering algorithm. Zhang *et al.* [13] has presented the Multi-Features Fusion image segmentation algorithm, which is used on high-resolution sensing images that contained more information about the spatial relationships between ground objects than low-resolution images. This algorithm was identifying the location of objects in the image that were being viewed. Huang *et al.* [14] were given two-way clustering based on the least crossing tree and DBSCAN calculations for picture division. Dubey *et al.* [15] presented the image segmentation-based clustering technique and several clustering strategies are discussed. Ramaraj and Niraimathi [16] has proposed using DBSCAN for real-time image segmentation based on super-pixels. Sudana *et al.* [17] has proposed the DBSCAN algorithm was employed to cluster images containing numerous intricate Balinese characters within a single large image.

Wang *et al.* [18] demonstrated using a new method to segment images very effectively based on the clustering algorithms being presented, primarily for use in applications involving images of the same size and shape as the image itself. Ramaraj and Niraimathi [19] presented the various clustering strategies that are utilized to characterize the organization of nanoscale assemblies in images obtained through localization microscopy. Said *et al.* [20] demonstrated a method for color-based image segmentation to divide colors into segments. The process of splitting up an image into distinct regions where each pixel has similar characteristics is known as image segmentation.

In study by Chena *et al.* [21] in this context, a novel color image segmentation algorithm called mean shift hierarchical clustering (MSHC) was introduced. Cong and Hiep [22] had presented a versatile and solo bunching approach in light of Voronoi locales, which could be applied to take care of the variety of image division issues. Shi *et al.* [23] a brand-new pixel intensity clustering algorithm for multi-level image segmentation was unveiled. Jinga *et al.* [24] also discussed the different clustering techniques for changing the standard fuzzy objective function by updating the membership and cluster centroids. The performance study of image segmentation techniques has been described by Khaleel *et al.* [25]. Bai *et al.* [26] various applications of the image segmentation problem in computer vision and image processing were presented. Fahrudin *et al.* [27] in recent years, the enhanced support vector clustering algorithm has gained significant interest for color image segmentation, particularly across various application fields. Reddy *et al.* [28] used the fuzzy clustering method for color image segmentation, producing a single image with uniform color.

## 3. METHOD

The pixel clustering heuristic determines the total number of clusters based on the expected pattern of related nodes using a new algorithm called fuzzy density-based spatial clustering of applications with noise (FDBSCAN). This approach aids in accurately identifying and grouping clusters within the data. The FDBSCAN clustering algorithm uses the notion of density reachability rather than mass or distance reachability to identify a cluster.

### 3.1. FDBSCAN clustering

The FDBSCAN clustering algorithm offers a sophisticated approach to predicting the age of tigers from camera trap images. FDBSCAN enhances the conventional DBSCAN approach by integrating fuzzy membership functions, which allow for a more nuanced classification of data points into clusters [29]. FDBSCAN operates by evaluating the density of data points within a specified radius, ϵ and midpoints, applying fuzzy logic to handle uncertainties and variations in the data. Unlike conventional clustering methods that assign data points to discrete clusters, FDBSCAN provides a degree of membership for each point, reflecting its potential association with multiple clusters [30]. This is particularly beneficial in handling the variability and complexity of natural images. Instead of strictly assigning each point to a single cluster or labeling it as noise, FDBSCAN assigns a degree of membership to each cluster. This flexibility helps manage the inherent uncertainty and variability in real-world datasets. Consequently, understanding how to choose the appropriate Epsilon and MinPoints values is critical. The FDBSCAN algorithms output is significantly altered by small changes in these values. MinPoints should have a value that is at least one higher than the dataset's number of dimensions.

$$MinPoints \geq Dimensions + 1 \tag{1}$$

Since this will make each point a distinct cluster, it makes no sense to take MinPoints as 1. In most cases, it needs to be at least 3. If the value of an epsilon chosen is too low, more clusters will be formed. And more data points considered to be noisy will become available for further analysis.

### 3.2. Reachability and connectivity

The FDSCAN clustering algorithm, when applied to the task of tiger age prediction, emphasizes the importance of reachability and connectivity within the data. Through these concepts, the algorithm effectively groups similar data points, ensuring that clusters formed are both cohesive and meaningful. This approach enhances the accuracy of predictions by ensuring that the algorithm can identify subtle patterns and relationships in the data, which are crucial for accurately estimating the age of tigers from camera trap images. Through the implementation of FDSCAN, the research achieves more reliable and precise clustering, directly contributing to the improved prediction of tiger ages in the field of wildlife research.

Here, if the distance $(X, Y) = \varepsilon$, $Y$ is a core point, then using $\varepsilon$ and MinPoints, point $X$ can be density-reachable directly from point $Y$ as Figure 1 illustrates. In this case, $X$ can be reached directly by density from $Y$, but the opposite is not true: A chain of points $p1, p2, p3, \ldots, pn$ and $p1 = X$ and $pn = Y$ makes $pi + 1$ directly density-reachable from $pi$, which means that a point $X$ is density-reachable from a point $Y$ in terms of epsilon, MinPoints. $N€(p) = q$ D dist $p\ q|(,)$ € defines the € (read as Eps)-neighborhood of a point $p$. It is possible to assert that $X$ is density-reachable from $Y$ because both $X$ and $Y$ are reachable from $O$; however, this is not the case for the other two dimensions. Figure 2 depicts the direct density-reachable color function. If there are two points, $p$ and $q$, the point $q$ is directly density-reachable from the other point as long as it is not more than a certain distance, like €; That is, $q$ is in the €-neighborhood of $p$. If $p1 = p$ and $pn = q$, then the point $q$ has a density that can be reached from $p$, and if $p$ is surrounded by many points, like $p1, p2, p3, \ldots pn$, then each point $pi + 1$ can be reached directly from $p$. For density-connected systems, as shown in Figure 3: density-reachable relationships are not symmetrical. The point q may be at the edge of a cluster because it lacks enough neighbors to be considered dense on its own. Two points p and q are density-connected if there is a point o that can be reached from both points p and q.

Figure 4 depicts the fuzzy DBSCAN-based clustering method used to determine the tiger's age. To eliminate the scale-dependency issue and obtain global values independent of the range of pixel data, the data are normalized for the parameters and MinPts.

$$MinPts = \varepsilon 2 . w^{max} \tag{2}$$

There are two coordinates between MinPts and ε values are fixed. For example, $w_i = |N(x_i;)|$, is where $Max_i = |1 \ldots n\, w_i . \varepsilon_1|$ in which $|N(x_i;\ \varepsilon_1)|$ denotes the set's fuzzy cardinality, $N(x_i;\ \varepsilon_1)$. The process's outcome may vary depending on the choice made by the function $Nx_i$, which may be any fuzzy neighborhood function. Likewise, $\varepsilon 2$ is obtained by (3).

$$\varepsilon 2 = \mu \frac{MinPts}{w^{max}} \tag{3}$$

For a constant $\varepsilon 1$, the fuzzy neighborhood set of point $x \in X$ is formed by (4):

$$FN(x, \varepsilon 1) = \{< y. N_x(y) > | y \in X, N_x(y) \geq \varepsilon 1 |\} \tag{4}$$

A fuzzy core point is obtained by (5):

$$card\ FuzzyN(x;\ \varepsilon 1, \varepsilon 2) \equiv \sum_{y \in N(x;\ \varepsilon 1)} Nx(y) \geq\ \varepsilon 2. \tag{5}$$

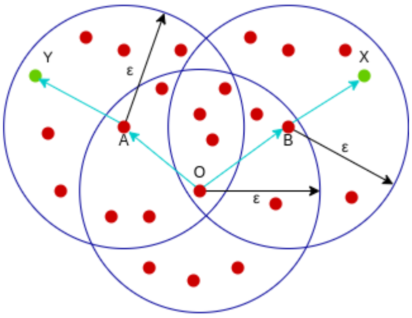


Figure 1. SplEps

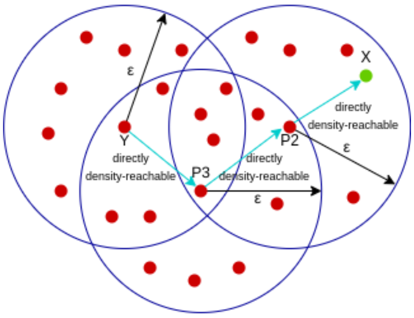


Figure 2. ColEps

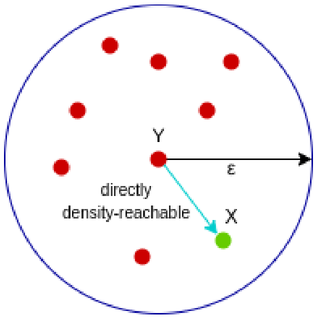


Figure 3. Connected

- Load the image dataset.
- Assign input parameters $\varepsilon 1$, $\varepsilon 2$ and $f$
- Initialize i = 0 and set the cluster assignment for all points as unassigned
- For each point p in the dataset:
    - Calculate cardinality of p by applying f to all points within distance epsilon p.
    - If cardinality (p) ≥ ε2
        - Assign p as a fuzzy core point
        - Start a new cluster $C_i$ and assign p to $C_i$
        - Find all points within ε1 distance of p and add to seed set S
        - For each point q in the S
            - Assign q to Ci
            - If cardinality q≥ ε2
                - Assign q as fuzzy core point
                - Find all points within ε1 distance of q and add to seed set S.
        - Set i = i + 1
- Mark all the points which remain unassigned as noise.
- End.

Figure 4. Fuzzy based DBSCAN clustering algorithm

As a result, the fuzzy DBSCAN method may be more resistant to the dataset's scale and density variations. The initial step in using fuzzy DBSCAN to determine a tiger's age is to load images. It has two parameters: MinPts, which indicates the density of points that serve as the core points, and MaxPts, which indicates the maximum distance between a data object and the image that defines neighboring points.

### 3.3. Computing points

FBDSCAN clustering algorithm, the process of computing points plays a vital fragment in accurately predicting the age of tigers. The algorithm determines the compactness of arguments within a specific radius $\epsilon$ and uses this to form clusters. The fuzzy membership function allows for a degree of uncertainty in the assignment of points to clusters, which is particularly useful in handling the inherent variability in tiger images. The core concept can be mathematically represented by the density function $f(x)$, where $(x)$ is a data point in the feature space:

$$f(x) = \sum_{i=1}^{n} exp\left(\frac{\|x-x_i\|^2}{2\sigma^2}\right) \tag{6}$$

Here, $n$ is the whole amount of points, $x_i$ represents the neighboring data points within the radius $\epsilon$ and $\sigma$ is a smoothing parameter that controls the spread of the Gaussian function. Once the density is computed, the algorithm applies fuzzy logic to assign membership values to each point, enabling it to handle noise and outliers more effectively. By optimizing these clusters through FBDSCAN, the algorithm enhances its ability to predict tiger age by identifying relevant patterns within the image data.

## 4. RESULTS AND DISCUSSION

The proposed model incorporates a tiger image database, making MATLAB tool usage easier. This database contains over 1,000 images from camera traps and other sources in different sizes and formats. Images of tigers from various age groups are all categorized within the same class, ensuring a consistent and efficient analysis process.

### 4.1. Computational complexity

When evaluated in terms of time erudition analysis to normalize their simulated efficiency, the poles apart cluster technique was found to be computationally complex and to be more efficient than the standard clustering technique. The ability to create hierarchies in bunching approaches was construed as a given computational convolution equation, whereas the fuzzy based DBSCAN clustering algorithm necessitates fewer steps to be performed in a given clustering method.

$$o((N - \sum_{t=0}^{m-1} N_r)^2) \tag{7}$$

Hence, $N$ represents the total number of color pixels, $m$ denotes the number of clusters, and $r$ indicates the number of iterations applied to $N$.

Figure 5 demonstrates the effect of spatial epsilon (EPS) and color EPS as parameters are determined. The image database is the focus of this efficient FDBSCAN clustering strategy. The sharp change in the k-dist value is similar to the correct Epsilon value. MinPts: The number of dimensions D in the data set can be used to calculate a minimum MinPts as follows: MinPts D+1. It makes no sense to set MinPts to a low value of 1, since then each point will already be a cluster on its own. With MinPts ≤ 2, The result will be equivalent to multiple levels of grouping using a single connection metric, with the dendrogram cut at level ε. ε: Using a k-distance graph, which plots the distance to the k=MinPts-1 nearest neighbor in order from the largest to the smallest value, the value of the distance can then be chosen.

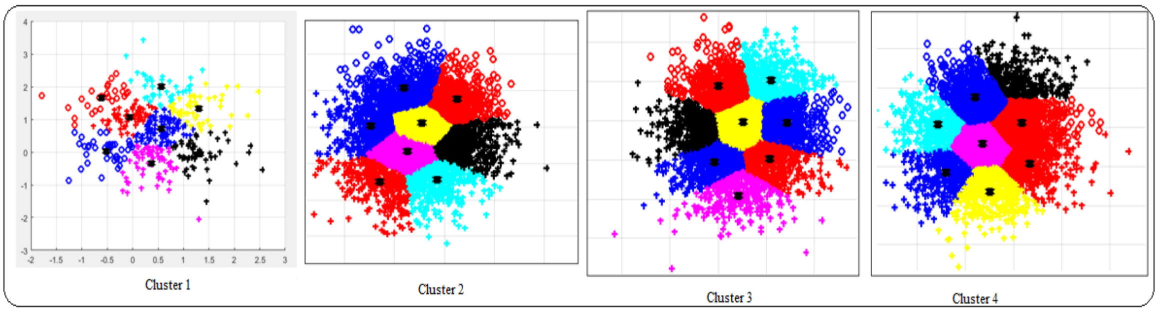


Figure 5. Pixel clustering based on tiger age group on k=2, k=4, k=6, plot values

### 4.2. Age prediction of the real time tiger image

Predicting the age of tigers based on their images involves segmenting the image into relevant regions and analyzing the pixel attributes to derive meaningful patterns. The underlying principle is that as tigers age, the characteristics of their fur, such as the density of stripes, color intensity, and texture, may change. The image of a tiger can be represented as a matrix of pixels, where each pixel has three color components (RGB). Let $I(x,y)$ represent the pixel intensity at coordinates $(x,y)$ in the image. The segmentation process can be expressed as a function that assigns each pixel to a cluster:

$$S(x,y) = \sum_{i=1}^{c} \mu_i(x,y).f_i\big(I(x,y)\big) \tag{8}$$

where $S(x,y)$ is the segmented output at pixel $(x,y)$. $\mu_i(x,y)$ is the membership value of pixel $(x,y)$ belonging to cluster $i$ (ranging from 0 to 1 in fuzzy clustering). $f_i\big(I(x,y)\big)$ is the function that represents the characteristics (like mean or variance) of cluster $i$. This difference represents the spatial separation needed between pixels handled by the cells and those in incorrect frames of $(T_{ij})$. Essentially, it measures how well the technique can distinguish between pixels in various frames. To assess overall segmentation precision, divide these values by the total number of pixels $(N = \sum R_i = \sum C_j)$,. This calculation involves determining $(\sum T_{ij})$, which represents the number of pixels in the tiger image database that align with the age classifications in the ground truth.

$$d = \frac{N \sum_{i=1,j=1,k=1}^{m} T_{ijk} - \sum_{i=1,j=1,k=1}^{m} R_i .C_j}{N^2 - \sum_{i=1,j=1,k=1}^{m} R_i .C_j} \tag{9}$$

For example, $d$ denotes the fundamental Euclidean distance, $N$ is the total number of pixels in an image, and m represents the number of red, green, blue (RGB) color classes. The total number of correctly classified pixels in a tiger image is indicated as $T_{ijk}$. Additionally, each color pixel's threshold value was set to a specific tiger when determining the age threshold for the tiger image. The number of pixels in each row and column is represented by $R_i$ and $C_j$ respectively.

The data above are sorted by year, as shown in Table 1. To compare with the different parameter like precision, recall, and F-measure is used to evaluate to predict the age of the tiger and find the efficient of the above metrics. The highest precision is 94.3 percent, while the lowest precision is 92%. The lowest recall is 91%, while the highest recall is 94%. The lowest F-measure is 91.6%, while the highest is 93.5%. The table's similarity measures are slightly excited when compared to each other. The most elevated accuracy is 94.3% for Euclidean and most noteworthy review is 94% for both comparability measures as city block and Chebyshev, and the most elevated F-measure is 93.5% in Euclidean and the least is found in the closeness measures on plain as to contrast and other one.

Table 1. Different parameters checked with 1 year tiger image data base

| Age | SMR | PV | RV | FMV |
|---|---|---|---|---|
| 1 year | CBV | 92% | 94% | 93% |
| | CCV | 93% | 94% | 93.5% |
| | EV | 94.3% | 91% | 92.65% |
| | MV | 92% | 91.2% | 91.6% |

Note: SMR-similarity measures, PV-precision value, RV-recall value, FMV: F-measures value

Additionally, Table 1 evaluates the consistency of clustering accuracy for different tiger image ages. It highlights specific functions such as city block, Chebyshev distance, Minkowski distance, and other distance measures, using clustering metrics like precision, recall, and F-measure. The experimental results are illustrated in Figure 6.

The following data are sorted by year in the tiger image database, as shown in Table 2. Consistently, the number of clusters is taken to be three. The highest precision recorded in the second year is 93%, while the lowest precision is 92%. The lowest recall is 90%, while the highest recall is 0.94. The lowest F-measure is 91.5%, while the highest is 93%. When compared to each of the similarity measures in the table, is slightly satisfied. The results are completely different from those of the previous year, despite the uniform change in the second year. City blocks have the highest precision at 93%, Minkowski has the highest recall at 94%, and Minkowski has the highest F-measure at 93.5%, while similarity measures on the tabular are the lowest when compared to other methods.

Similarity-based clustering can effectively predict the age of tiger images, as demonstrated by the detailed metrics for city block, Chebyshev, Minkowski, and Euclidean distances in Table 2. Accuracy is assessed using clustering indicators such as fit rate, recall, and F-value. The results of these experiments are illustrated in Figure 7.

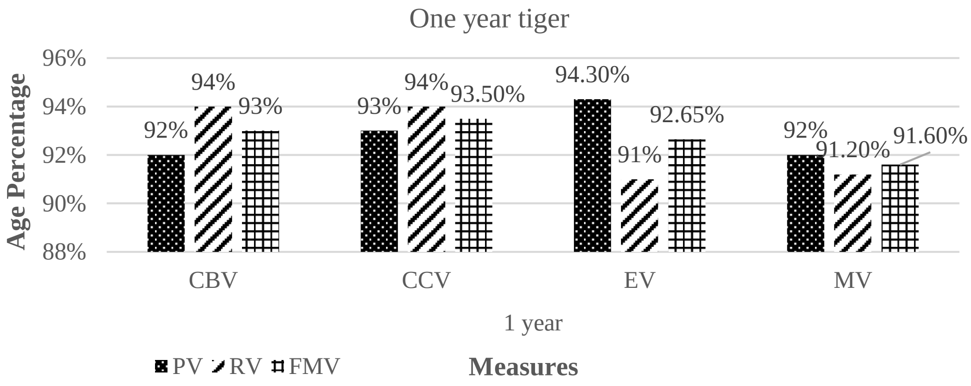


Figure 6. Numerous similarity metrics are used on FDBSCAN with one-year tiger image

Table 2. Applied FDBSCAN with different similarity measures using two-year tiger image

| Age | SMR | PV | RV | FMV |
|---|---|---|---|---|
| 2 years | CBV | 93% | 90% | 91.5% |
| | CCV | 91% | 92% | 91.5% |
| | EV | 92% | 93% | 92.5% |
| | MV | 92% | 94% | 93% |

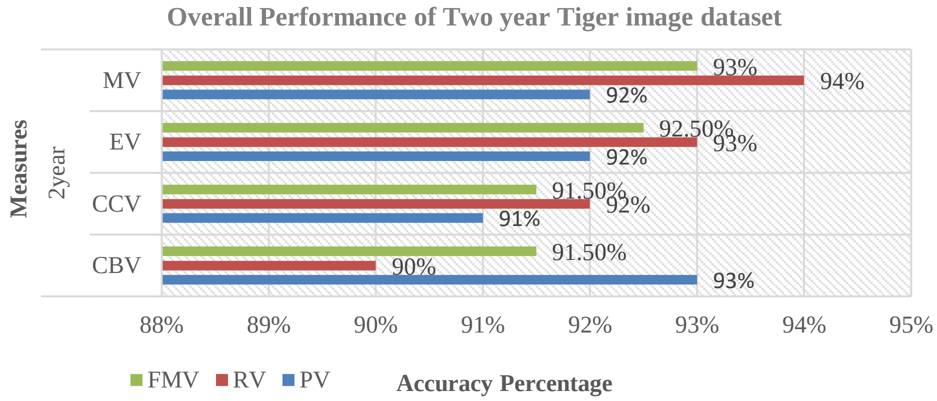


Figure 7. Different similarity metrics are used on FDBSCAN with 2nd -years tiger image

As per Table 3, the information is arranged year wise. Over the course of 15 years, the lowest precision was 92%, while the highest precision was 93%. The lowest recall rate is 90.10%, while the highest recall is 95.5%. The lowest and highest F-measures are 91.05% and 94%, respectively. When compared to each of the similarity measures in the table, is slightly satisfied. The 15th year sees a consistent change, but the outcomes are completely different from those of the previous year. The city block has the highest precision of 93%, the highest recall of 95%, the highest F-measure of 94%, and the lowest similarity measures on the tabular to compare with other ones.

Table 4 presents the consistency of various clustering matrices, including root mean square error (RMSE) values, prediction time intervals, and image analysis time, for both established and improved methods. The clustering results are presented graphically whenever the proposed algorithms were utilized and produced superior results that were significantly more accurate and efficient. In fuzzy based DBSCAN clustering, the proposed method has the highest accuracy rating of any of the three methods tested.

Table 3. FDBSCAN used different clustering similarity approaches on tiger image dataset

| Age | SMR | PV | RV | FMV |
|---|---|---|---|---|
| 15 years | CBV | 93% | 95% | 94% |
| | CCV | 92% | 91% | 91.5% |
| | EV | 93% | 93% | 93% |
| | MV | 92% | 90.10% | 91.05% |

Table 4. Evaluation of enhanced clustering metrics

| Accuracy value | | RMSE value | | Time process | | Image retrieval process | |
|---|---|---|---|---|---|---|---|
| Prevailing | 84.489 | Prevailing | 0.765889 | Prevailing | 3.27492 | Prevailing | 2.51 |
| Enhanced | 85.904 | Enhanced | 0.678467 | Enhanced | 2.456657 | Enhanced | 1.86 |

Table 3 measures the accuracy of similarity-based clustering, using similarity functions such as city block, Chebyshev distance, Minkowski distance, and Euclidean distance. Clustering indicators like precision, recall, and the F-measure are employed to evaluate these functions' effectiveness in predicting the age of tiger images. The results of these experiments are illustrated in Figure 8. Figure 9 compares the accuracy, processing time, and image search performance of the tiger image database with both the proposed and existing methods. Table 4 details the results of the proposed approach.

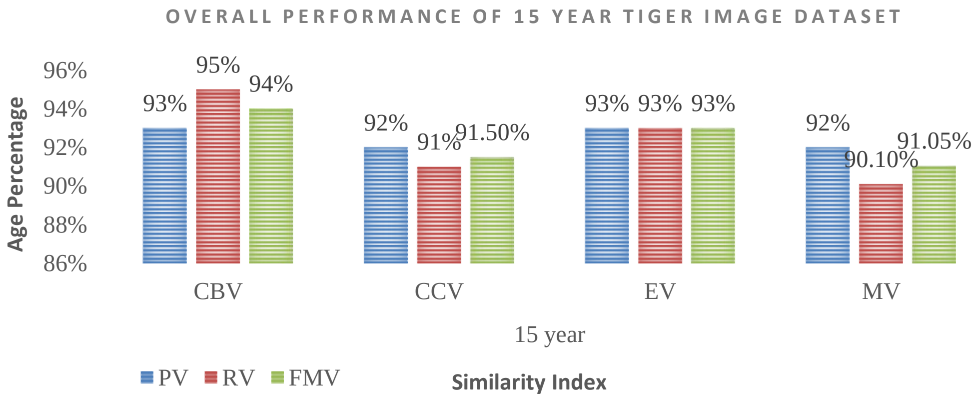


Figure 8. Illustrate on FDBSCAN used with different clustering similarity methods applied the tiger image

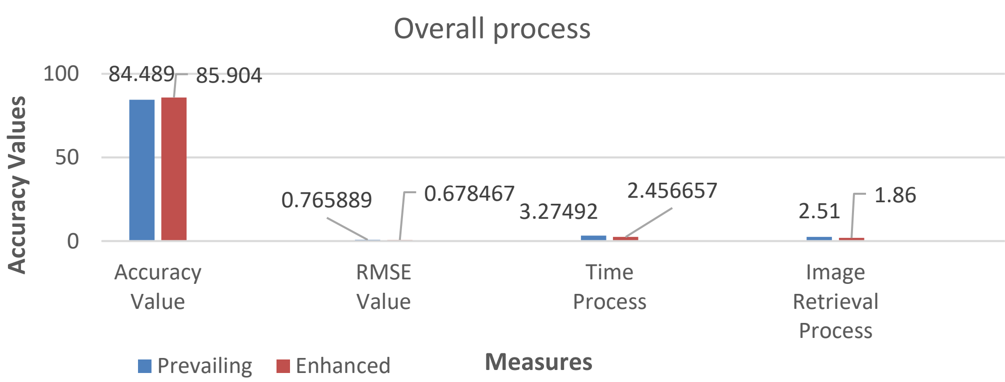


Figure 9. Comparison of overall performance measures

## 5. CONCLUSION

In conclusion, this research introduces a sophisticated approach utilizing fuzzy clustering models, particularly FDBSCAN, to enhance the accuracy and efficiency of predicting tiger age based on color attributes from camera trap images. Compared to the latest statistical methods, our approach excels in both precision and speed, offering a significant improvement in the visual presentation and analysis of image databases. The primary objective of automating tiger age determination through image databases has been successfully achieved by processing over 1,000 real-time images of tigers in their natural forest habitats. This extensive dataset includes a diverse range of adult specimens and is analyzed based on distinctive features such as stripes, skin tones, and overall color variations. Through classifying images into different age groups, when the method effectively groups tigers according to their age. The research is structured into detailed sections that elucidate the correlation between color characteristics and age determination, showcasing the method's robustness and applicability. This innovation not only advances the field of image processing but also provides valuable insights for wildlife research and real-world applications, offering a significant contribution to both theoretical and practical aspects of age prediction in tigers.

## REFERENCES

[1] K. Gautam and R. Singhai, "Color image segmentation using particle swarm optimization in lab color space," *International Journal of Engineering Development and Research*, vol. 6, no. 1, pp. 373–377, 2018.

[2] E. Huang *et al.*, "Occlusion-resistant instance segmentation of piglets in farrowing pens using center clustering network," *Computers and Electronics in Agriculture*, vol. 210, Jul. 2023, doi: 10.1016/j.compag.2023.107950.

[3] X. Liu, C. Hu, and P. Li, "Automatic segmentation of overlapped poplar seedling leaves combining Mask R-CNN and DBSCAN," *Computers and Electronics in Agriculture*, vol. 178, 2020, doi: 10.1016/j.compag.2020.105753.

[4] T. Czerniawski, B. Sankaran, M. Nahangi, C. Haas, and F. Leite, "6D DBSCAN-based segmentation of building point clouds for planar object classification," *Automation in Construction*, vol. 88, pp. 44–58, 2018, doi: 10.1016/j.autcon.2017.12.029.

[5] K. Mardani and K. Maghooli, "Enhancing retinal blood vessel segmentation in medical images using combined segmentation modes extracted by DBSCAN and morphological reconstruction," *Biomedical Signal Processing and Control*, vol. 69, 2021, doi: 10.1016/j.bspc.2021.102837.

[6] N. Hanafi and H. Saadatfar, "A fast DBSCAN algorithm for big data based on efficient density calculation," *Expert Systems with Applications*, vol. 203, 2022, doi: 10.1016/j.eswa.2022.117501.

[7] X. Huang, T. Ma, C. Liu, and S. Liu, "GriT-DBSCAN: a spatial clustering algorithm for very large databases," *Pattern Recognition*, vol. 142, 2023, doi: 10.1016/j.patcog.2023.109658.

[8] S. Yang, P. Li, H. X. Wen, Y. Xie, and Z. He, "K-hyperline clustering-based color image segmentation robust to illumination changes," *Symmetry*, vol. 10, no. 11, 2018, doi: 10.3390/sym10110610.

[9] M. Ramaraj and S. Niraimathi, "Comparative analysis of proposed methods for analysing color pixel based image segmentation using tiger image dataset," *International Journal of Research*, vol. 7, no. 9, pp. 245–252, 2018.

[10] Q. Zhu, X. Tang, and A. Elahi, "Application of the novel harmony search optimization algorithm for DBSCAN clustering," *Expert Systems with Applications*, vol. 178, 2021, doi: 10.1016/j.eswa.2021.115054.

[11] M. Yan, Y. Chen, Y. Chen, G. Zeng, X. Hu, and J. Du, "A lightweight weakly supervised learning segmentation algorithm for imbalanced image based on rotation density peaks," *Knowledge-Based Systems*, vol. 244, 2022, doi: 10.1016/j.knosys.2022.108513.

[12] C. Chen, J. Guo, Y. Li, and L. Xu, "Segmentation-based hierarchical interpolation filter using both geometric and radiometric features for LiDAR point clouds over complex scenarios," *Measurement: Journal of the International Measurement Confederation*, vol. 211, 2023, doi: 10.1016/j.measurement.2023.112668.

[13] Y. Zhang, Y. Zhang, L. Fan, and N. Wang, "Fast and accurate superpixel segmentation algorithm with a guidance image," *Image and Vision Computing*, vol. 129, 2023, doi: 10.1016/j.imavis.2022.104596.

[14] Y. Huang, W. J. Huang, X. L. Xiang, and J. J. Yan, "An empirical study of personalized advertising recommendation based on DBSCAN clustering of sina weibo user-generated content," *Procedia Computer Science*, vol. 183, pp. 303–310, 2021, doi: 10.1016/j.procs.2021.02.063.

[15] S. K. Dubey and S. Vijay, "A review of image segmentation using clustering methods," *International Journal of Applied Engineering Research*, vol. 13, no. 5, pp. 2484–2489, 2018.

[16] M. Ramaraj and S. Niraimathi, "Enhanced image segmentation to infer the age of the tiger using fuzzy modified K-means clustering algorithm (FMKMCA)," *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, vol. 10, no. 2, 2019.

[17] O. Sudana, I. K. G. D. Putra, M. Sudarma, R. S. Hartati, and A. Wirdiani, "Image clustering of complex balinese character with DBSCAN algorithm," *Journal of Engineering Technology*, vol. 6, no. 1, pp. 548–558, 2018.

[18] Y. Wang, X. Yang, L. Zhang, X. Fan, Q. Ye, and L. Fu, "Individual tree segmentation and tree-counting using supervised clustering," *Computers and Electronics in Agriculture*, vol. 205, 2023, doi: 10.1016/j.compag.2023.107629.

[19] M. Ramaraj and S. Niraimathi, "Design and implementations of color pixel based image segmentation using enhanced data clustering algorithms to applying on tiger image dataset," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 18, pp. 2719–2739, 2018.

[20] N. S. B. Mat Said, H. Madzin, S. K. Ali, and N. S. Beng, "Comparison of color-based feature extraction methods in banana leaf diseases classification using SVM and K-NN," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 3, pp. 1523–1533, 2021, doi: 10.11591/ijeecs.v24.i3.pp1523-1533.

[21] H. Chen, M. Liang, W. Liu, W. Wang, and P. X. Liu, "An approach to boundary detection for 3D point clouds based on DBSCAN clustering," *Pattern Recognition*, vol. 124, 2022, doi: 10.1016/j.patcog.2021.108431.

[22] V. D. Cong and T. T. Hiep, "Support vector machine-based object classification for robot arm system," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 5, pp. 5047–5053, 2023, doi: 10.11591/ijece.v13i5.pp5047-5053.

[23] P. Shi, L. Guo, H. Cui, and L. Chen, "Geometric consistent fuzzy cluster ensemble with membership reconstruction for image segmentation," *Digital Signal Processing: A Review Journal*, vol. 134, 2023, doi: 10.1016/j.dsp.2022.103901.

[24] W. Jing, C. Zhao, and C. Jiang, "An improvement method of DBSCAN algorithm on cloud computing," *Procedia Computer Science*, vol. 147, pp. 596–604, 2019, doi: 10.1016/j.procs.2019.01.208.

[25] A. D. Khaleel *et al.*, "Pattern reconfigurable dielectric resonator antenna using capacitor loading for internet of things applications," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 4, pp. 4145–4151, 2023, doi: 10.11591/ijece.v13i4.pp4145-4151.

[26] X. Bai, Z. Xie, X. Xu, and Y. Xiao, "An adaptive threshold fast DBSCAN algorithm with preserved trajectory feature points for vessel trajectory clustering," *Ocean Engineering*, vol. 280, 2023, doi: 10.1016/j.oceaneng.2023.114930.

[27] A. E. Fahrudin, Endarko, K. Ain, and A. Rubiyanto, "Enhanced image reconstruction of electrical impedance tomography using simultaneous algebraic reconstruction technique and K-means clustering," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 4, pp. 3987–3997, 2023, doi: 10.11591/ijece.v13i4.pp3987-3997.

[28] S. S. Reddy, V. V. R. M. Rao, K. Sravani, and S. Nrusimhadri, "Image quality evaluation: evaluation of the image quality of actual images by using machine learning models," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 2, pp. 1172–1182, 2024, doi: 10.11591/eei.v13i2.5947.

[29] A. Ashraf, A. Sophian, A. A. Shafie, T. S. Gunawan, and N. N. Ismail, "Machine learning-based pavement crack detection, classification, and characterization: a review," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 6, pp. 3601–3619, 2023, doi: 10.11591/eei.v12i6.5345.

[30] E. L. Cahapin, B. A. Malabag, C. S. Santiago, J. L. Reyes, G. S. Legaspi, and K. L. Adrales, "Clustering of students admission data using k-means, hierarchical, and DBSCAN algorithms," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 6, pp. 3647–3656, 2023, doi: 10.11591/eei.v12i6.4849.


## BIOGRAPHIES OF AUTHORS

**Ramaraj Muniappan** is working as an assistant professor in the Department of Computer Science at Rathinam College of Arts and Science, Coimbatore. He holds a Ph.D. degree in computer science from Bharathiar University in 2020 with a specialization in data mining with image processing and fuzzy logic in image analysis. His research areas are data mining, image processing, fuzzy logic, pattern recognition, and deep learning concepts. He has published numerous research articles in reputed national and international journals and has also filed patents in the same field. He is a reviewer for many international journals, including IEEE, ASTESJ, and JERS. He can be contacted at email: ramaraj.phdcs@gmail.com, ramaraj.phdcs123@gmail.com.

**Srividhya Selvaraj** is working as an associate professor in the department of Computer Science at KPR College of Arts Science a Research, Coimbatore. She has more than a decade of teaching experience in different verticals. She is a reviewer in various international journals and life-time member in IAENG. Her research area includes data mining, machine learning and deep learning. She has about 10+ research papers published in national and international journals. She also presented more than 15 papers in national and international conferences and also published many books. She can be contacted at email: vidhyasai14@gmail.com.

**Rani Vanathi Gurusamy** is currently working as a professor and head at Sri Ramakrishna CAS for Women. She has been associated with the same college for the last 24 years. She received her Ph.D. in computer science from Bharathiar University in 2013. Her areas of research include data science, mobile ad hoc networks, cloud computing, network security, and online education. She has published more than 30 papers in national and international journals, including SCOPUS, Thomson Reuters, UGC, and IEEE, and has also presented at various national and international conferences. Her latest book publication is on the topic of multimedia systems. She received ₹6,00,000 funding assistance from ICSSR IMPRESS in 2019. She was awarded the "Research Award" by Bharathiar University on National Science Day in 2021. She is the current vice-president of the SRCW alumni association and the research coordinator of SRCW. She can be contacted at email: ranics@srcw.ac.in.

**Velumani Thiyagarajan** is working as an assistant professor in Department of Computer Science at Rathinam College of Arts and Science (Autonomous), Coimbatore 641 021 from 14-09-2020 to till date. He has got more than 12 years of teaching experience. He has obtained his B.Sc. (CT), M.Sc. (CS), M.Phil and MBA degrees from Periyar University Salem, Tamil Nadu, India. He has obtained his B.Ed. Degree from Indira Gandhi National Open University (IGNOU) at Delhi, his M.Sc. (Psychology) from Madras University Madras, Tamil Nadu, India. He has obtained his Ph.D. in Computer Science from Manonmaniam Sundaranar University at Tirunelvelli, Tamil Nadu, India. His area of interest is image processing research directions involved knowledge discovery and data mining, pattern recognition, knowledge-based neural networks, software engineering and IoT. He has published more than 15 papers in topmost international peer-reviewed journals and conference. He can be contacted at email: velumani46@gmail.com.

**Dhendapani Sabareeswaran** is working as an assistant professor in the Department of Computer Science at LRG college for women, Tiruppur. He holds a Ph.D., degree in computer science at Karpagam Academy of Higher Education in the year of 2020 with specalization in data mining. His research areas are data mining, image processing, fuzzy logic, pattern recognition. He has published more research article in the reputed various national and international journals and also filed the patents in the same field. He can be contacted at email: sabaredhandapani@gmail.com.

**David Prasanth** is working as an assistant professor in the Department of Computer Science at Shri Nehru Maha Vidyalaya College of Arts and Science, Coimbatore, Tamil Nadu in India. He received his bachelor's degree and master's degree from Bharathiar University Coimbatore in 2022 and 2024 respectively. His research areas are data mining, in the various fields. He has published more research article in the reputed various national and international journals and also filed the patents in the same field. He can be contacted at email: prasanthedisonon17@gmail.com.

**Varadharaj Krithika** is working as an assistant professor in the Department of Computer Science and Engineering at Akshaya College of Engineering and Technology, Coimbatore, Tamil Nadu in India. She received her bachelor's degree and master's degree from Anna University Chennai in 2013 and 2015 respectively. She is currently pursuing her Ph.D. She has 7+ years of teaching experience. She is a lifetime member of Indian Society of Technical Education (ISTE). Her area of interest is image processing and deep learning. He can be contacted at email: krithikav@acetcbe.edu.in.

**Bhaarathi Ilango** is working as an assistant professor in the Department of Computer Science at Rathinam College of Arts and Science, Coimbatore. He holds a NET. Qualification in computer science at NTA in the year of 2023. with specialization in data mining with image process and also fuzzy logic in the image analysis. His research areas are data mining, in the various fields. He has published more research article in the reputed various national and international journals and also filed the patents in the same field. He can be contacted at email: bhaarathi.cs@rathinam.in.

**Dhinakaran Subramanian** MCA., M.Phil., Ph.D., controller of examinations and associate professor in computer science, Rathinam College of Arts and Science (Autonomous), Coimbatore – 641021, Tamil Nadu, India. He has sixteen years of dedicated teaching experience to excel in his field of research work. He has written several research articles in international, peer reviewed and Scopus indexed journals. He published in reputed national and international journals in different pasture of research. He has presented papers at various national level and international level conferences. He is also acting as subject expert of the board of studies for various leading autonomous colleges. He has written in six books. He is also acting as a subject expert of the board of studies for Bharathiar University. He is a former senate member of Bharathiar University. He can be contacted at: dhinakaran.cs@gmail.com.