IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol. 9, Issue.6, June 2021 E-ISSN: 2347-2693

Image Color Segmentation With Kdtree Library For Car Color Identity Classification

Joni[1]*, Erwin[2]

[1] Computer Science Dept, STMIK TIME, Medan, Indonesia
[2] Accounting Dept, STIE Mikroskil, Medan, Indonesia

*Corresponding Author: joni_hgw@yahoo.com

DOI: https://doi.org/10.26438/ijcse/v9i6.912 | Available online at: www.ijcseonline.org

Received: 03/Jun/2021, Accepted: 13/Jun/2021, Published: 30/Jun/2021

***Abstract*—** Joni and Erwin, *Abstract*, Artificial Intelligence (AI) has been widely used in analyzing objects, such as text, image, etc. There are many things that can be analyzed from images for the needs of identifying and classifying objects into certain types. One of the identifiable data is color. To identify the main color of an object, a vehicle image (car) requires a very complex analysis process. In this study, the identification process was carried out using an image center area analysis approach. This is based on the perception that the main color is in the middle of the object area. All color pixels in the analyzed area are converted to color names using the KDTree library. The segmentation process will produce several groups of color values. From the color matrix that has been through the segmentation process, the color identity of the object is obtained, which is determined by the mode value of the color matrix.

***Keywords*—** Color Segmentation, Color Identification, KDTree Library, Car Color

I. INTRODUCTION

Computer technology that is growing rapidly, especially in the field of Artificial Intelligence (AI) has been widely used in analyzing objects and phenomena that occur. One of the implementations of this scientific field is the identification and classification of object identities. The characteristics of an object can be represented through images.

By analyzing image and utilizing AI, some difficulties in various aspects of life can be overcome. An example of the implementation of the development of this science in the medical field is the detection of symptoms of heart cancer [1]. Another example in the culinary field is the classification and identification of types of food [2].
The process of recognizing and determining the characteristics of an image object is quite difficult to do. In this process, by utilizing the technology developed in AI, it will be easier to take a role in the classification of 2-dimensional image objects.

In determining the identity or classifying objects into certain types, it is necessary to process the separation of parts (segments) of the image. This separation process is known as the segmentation process. Through segmentation, digital data will be obtained from the analyzed image object. There are many things that can be analyzed from images for the needs of identifying and classifying objects into certain types. One of the identifiable data is color, because color is the main characteristic of the object.

Vehicles are one of the many objects identified by their main color. In this study, the color characteristics of the vehicle image will be analyzed, in which case the type of vehicle is a car. Car color is one of the main identities of the car. The importance of the identity of the car's color is regulated in the Regulation of the Head of the National Police of the Republic of Indonesia Number 5 of 2012 concerning Registration and Identification of Motor Vehicles. Article 37 paragraph 1 explains that one of the data contained in the STNK (Vehicle Registration Certificate) is color.

A car can be classified into a certain type of color based on its main color. The object of the car, in this study, is represented by a 2-dimensional image of the car. The car image analyzed is a presentation of the side view image display. Because there are many color groups in the car image, in this study the center area of the image is analyzed. The selection of this area is based on the assumption that the main color of the object is in the middle area of the image that presents the object. With this assumption, the main color in the center area of the image is stated as the main color of the car.

II. LITERATURE REVIEW

The segmentation process is one of the keys in getting an object recognition or detection result. Segmentation divides an image into parts or segments that are simpler and more meaningful so that further analysis can be carried out. Typically image segmentation is the process used to locate objects and boundaries (e.g., lines or curves) in images. Furthermore, it can be defined as the process of

© 2021, IJCSE All Rights Reserved

labeling every pixel in an image, where all pixels having the same label share certain visual characteristics [3].

To perform the segmentation process of the image object, various approaches can be determined. In general, the image segmentation approach that is often used is the intensity approach, the color approach and the shape approach. Usually segmentation uses local information in the digital image to compute the best segmentation, such as color information used to create histograms or information indicating edges, boundaries, or texture information [4].

Color segmentation is the separation of parts in an image based on the colors contained in the image. During its development, various methods have been carried out for color segmentation, such as the clustering method and the index method.

Color image segmentation that is based on the color feature of image pixels assumes that homogeneous colors in the image correspond to separate clusters and hence meaningful objects in the image. In other words, each cluster defines a class of pixels that share similar color properties. As the segmentation results depend on the used color space, there is no single color space that can provide acceptable results for all kinds of images. For this reason, many authors tried to determine the color space that will suit their specific color image segmentation problem [5].
The smallest part or unit of the image object is the pixel. Each pixel presents certain digital information. In analyzing images, digital data can be operated mathematically. This process of image object pixel analysis, in relation to image segmentation, will be completed when the object being analyzed in the application is isolated.

In general, segmentation algorithms are based on one of two basic properties of intensity values:
1) Discontinuous: divides an image based on changes in the intensity value (for example, sides)
2) Similarity: dividing an image based on similarity according to certain defined criteria.

III. RELATED WORK

In research [6], experiments were conducted on the color image segmentation process on microscopic images of tuberculosis bacteria (mycobacterium tuberculosis) derived from patient sputum, as color image samples that have a very complex variation in intensity values. Sputum obtained from patients was stained with the Ziehl-Neelsen staining method. The results of this staining provide complex slide images, due to the different staining results depending on the skills of the laboratory staff, so that clinical officers have difficulty when performing manual slide examinations. To assist clinical officers in reading slides, color segmentation of the slide image was carried out to extract the TB bacteria image and remove the background image.

Several methods have been carried out in this research, namely adaptive color thresholding in the RGB, HSV, CIE L*a*b color space, which gives good segmentation results in the CIE L*a*b color space. Then tried the K-Means Clustering and K-Nearest Neighbors segmentation methods to improve the adaptive color thresholding color segmentation performance, and the K-Nearest Neighbors method.

In the study to identify the quality of rice, the color segmentation of the rice image was carried out [7]. Good quality rice which has a white calcified color and damaged quality rice which has a yellowish color. To determine the difference in the color of good quality and damaged rice, identification errors often occur due to differences in color perception. This can be overcome by creating a system to identify good quality and damaged IR 64 rice grains. The data used is primary data, in the form of images of good quality rice grains as many as 50 and 50 grains of damaged rice. The estimator variables tested were La*b* and Sa*b* using the Adaptive Neuro Fuzzy Inference Systems (ANFIS) method.

IV. METHODOLOGY

In this study, the car object to be analyzed is represented by presenting a 2-dimensional image with a side view, as shown in Figure 1.

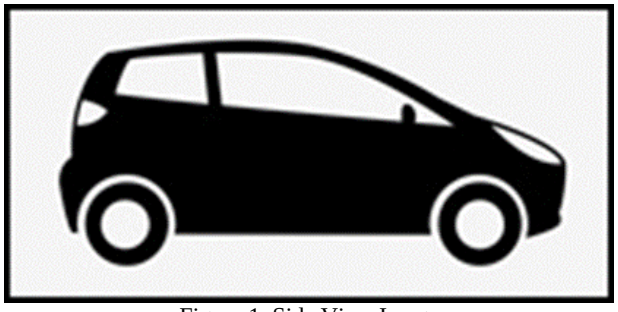


Figure 1. Side View Image

The analysis is carried out on the image center area. The center area is defined as 25% in the center of the image. Each pixel in the middle area is analyzed to obtain digital data which is the color of each pixel. The area in question is shown in Figure 2.

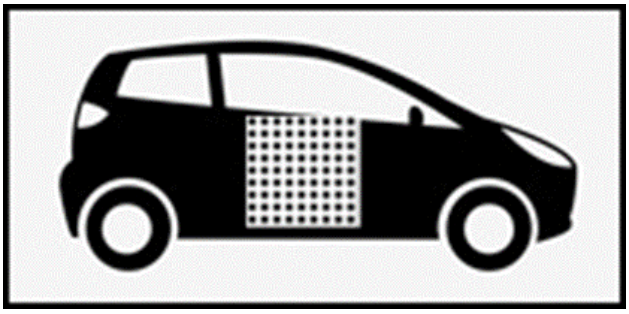


Figure 2. Image Center Area of Car

Determination of the center area in this image is calculated as follows:

© 2021, IJCSE All Rights Reserved

i=(image_height-25%×image_height)/2 ….. (1)
j=(image_width-25%×image_width)/2 ….. (2)

Next, from equations (1) and (2) an image pixel matrix is formed as follows:

$$A_{m\times n}=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

where,
a_11=a_ij ……..... (3)
m=j+25%×image_width ……..... (4)
n=i+25%×image_height ……..... (5)

Each matrix element represents each pixel in the center area of the image. Next, image segmentation is carried out to produce color information from the pixels it represents, in the form of RGB (Red, Green, Blue).

By implementing the KDTree library, color information is in RGB form, converted to color names (in English), via the syntax:

```
color_name = KDTree (color_RGB)
```

The converted color information is then arranged in a color table complete with the frequency of appearance of each color, as shown in Table 1.

Table 1. Image Center Area Color Name List

No.	Color Name	Frequency
1	color_name $_1$	num $_1$
2	color_name $_2$	num $_2$
…	…	…
3	color_name $_n$	num $_n$

The color value that appears most frequently (mode value) from a predefined list of colors is the color name with the largest number number. The largest number of figures is:

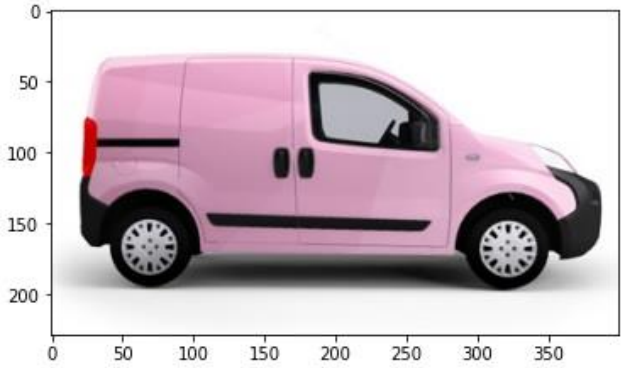
```
max_num = max (Frequency)
```

V. RESULTS AND DISCUSSION

The phase of identifying the color of the car image are described through the following steps:

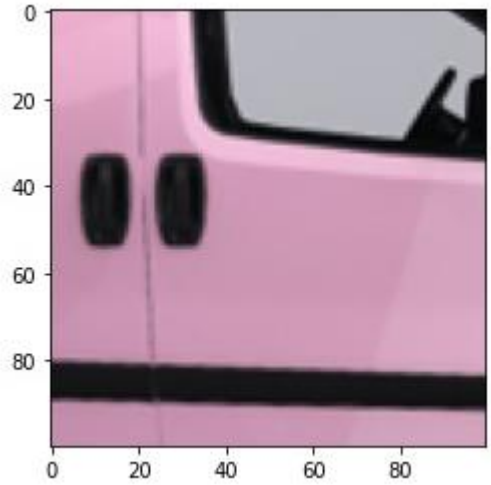
1) The image of the object which will be analyzed is scanned and presented in the form of a plot using the matplotlib library. Through this process, the required information from the image can be obtained as a research study material.

```
<matplotlib.image.AxesImage at 0x10d4a1a8>
```



2) To obtain the center area of the image, cropping is done, so that the image area is obtained as a color segmentation target. The center area of the image is 25% of the total area of the intact image.

```
Img-Width :  400
Img-Height :  229
Left-Margin :  150
Right-Margin :  250
Top-Margin :  64
Bottom-Margin :  164
```



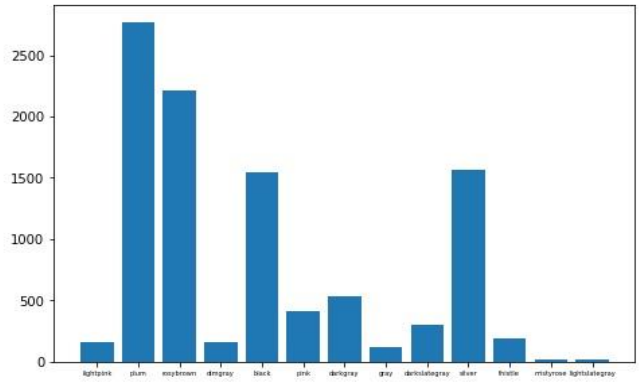
3) Next, color segmentation is carried out in the middle area of the image by defining the color for each image pixel. Color definitions are categorized in RGB color mode. Each pixel is associated as a matrix element, where each element has 3 values (R, G, B).

No.	Red	Green	Blue
1	238	175	206
2	238	175	206
3	237	174	205
4	236	173	204
5	234	171	202
..	...	...	...

4) Using the KDTree library, the color values in RGB mode are then converted to color names. The same colors are grouped so that the number of pixels of each color can be known. The list of colors that have

© 2021, IJCSE All Rights Reserved

been generated by the segmentation process is presented in the form of a bar chart. From the presentation of this data, a color classification will be obtained that shows the color identity of the object. The color with the highest number is the main color of the object.



5) Color lists can also be presented in tabular form. From this information, the color that appears most frequently is the color with the highest amount (intensity). This color is the main color of the object, so it can be said as the color identity of the object.

No.	Color	Count
1	lightpink	156
2	plum	2774
3	rosybrown	2209
4	dimgray	162
5	black	1543
6	pink	412
7	darkgray	539
8	gray	117
9	darkslategray	299
10	silver	1568
11	thistle	186
12	mistyrose	21
13	lightslategray	14

Color Identity : plum

VI. CONCLUSION AND FUTURE SCOPE

By applying image segmentation, color tabulation of the object can be obtained, so that the main color of the object can be further identified. The way of presenting the image display greatly determines the success of the object color analysis method with the image center area approach. By presenting an image with a side view, image segmentation in the middle area of the image is sufficient to represent the color of the object. Classifying object colors by this method will only produce limited color types (names) that are registered in the KDTree library.

The experimental result shows that the color identification which is the color identity of the object depends on the color segmentation in the center area of the image. If the center area of the image is an area that displays the main color of the car, the accuracy of determining the color identity of the car will be higher.

This research can be developed so that color identification is not only done through the analysis of certain areas of the object, but covers the entire object area. Before segmentation is done, it can be started by detecting the position of the object in the image.

REFERENCES

[1] Astha Pathak, Avinash Dhole, "Image classification Method in detecting Lungs Cancer using CT images: A Review," International Journal of Computer Sciences and Engineering, **Vol.9, Issue.5, pp.37-42, 2021.**

[2] Yash Baid, Avinash Dhole, "Food Image Classification Using Machine Learning Techniques: A Review," International Journal of Computer Sciences and Engineering, **Vol.9, Issue.5, pp.31-36, 2021.**

[3] M. Lalitha, M. Kiruthiga, and C. Loganathan, "A survey on image segmentation through clustering algorithm," International Journal of Science and Research, **vol. 2, no. 2, pp. 348–358, 2013.**

[4] N. Sharma, M. Mishra, and M. Shrivastava, "Colour image segmentaion techniques and issues: an approach," International Journal of Scientific & Technology Research, **vol. 1, no. 4, pp. 9–12, 2012.**

[5] L. Busin, N. Vandenbroucke, and L. Macaire, "Color spaces and image segmentation," Advances in Imaging and Electron Physics, **vol. 151, pp. 65–168, 2008.**

[6] R. Rulaningtyas, A. B. Suksmono, T. L. R. Mengko, G. A. P. Saptawati, "Segmentasi Citra Berwarna dengan Menggunakan Metode Clustering Berbasis Patch untuk Identifikasi Mycobacterium Tuberculosis" (Color Image Segmentation Using Patch-Based Clustering Method for Identification of Mycobacterium Tuberculosis), Jurnal Biosains Pascasarjana, **Vol. 17, No. 1, pp. 19-25, August 2015.**

[7] M. G. Alfianto, R. N. Whidhiasih, Maimunah, "Identifikasi Beras Berdasarkan Warna Menggunakan Adaptive Neuro Fuzzy Inference System" (Rice Identification Based on Color Using Adaptive Neuro Fuzzy Inference System), Jurnal Penelitian Ilmu Komputer, Sistem Embedded & Logic, **Vol. 5, No. 2, pp. 51-59, 2017.**

AUTHORS PROFILE

Joni received his Bachelor of Computer Science from Catholic University of St. Thomas, Medan, Indonesia. He has about 10 years of experience in teaching and research. At present, he is teaching Computer Science in STMIK TIME, a Computer Science College which is located in Medan, North Sumatera, Indonesia.

Erwin received his Bachelor of Math Science from University of North Sumatera, Medan, Indonesia. He has about 10 years of experience in teaching and research. At present, he is teaching Math and Statistics in STIE Mikroskil, an Economic College which is located in Medan, North Sumatera, Indonesia.

© 2021, IJCSE All Rights Reserved