International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 2309-2315 www.ijaem.net ISSN: 2395-5252

Color Detection System for Diamond Sorting Using Machine Learning

Mary Echabuo Afor[1], Tarcisius Forjong Dewingong[2], Bello Ibrahim Aliyu[3] Pradeep Kumar Mishra[4], and Gouri Sankar Mishra[5]

[1,2,3,4,5]*Department of Computer Science and Engineering, School of Engineering & Technology, Sharda University, Greater Noida, India.*

Submitted: 20-05-2022 Revised: 28-05-2022 Accepted: 30-05-2022

ABSTRACT : This paper presents a concept of color detection for diamond sorting. Using machine learning to teach the system how to recognize color in RGB space and predict its name and values. The program gathers images from the dataset and analyses the diamond's color attributes. Diamond classification is based on models created with the aid of machine learning techniques and image analysis algorithms. OpenCV is used for computer vision and to process the picture that has to be identified, while Pandas, a Python library, collects and analyses the data from the dataset to provide a suitable forecast of the color name and values. The Color Detection technique is implemented in image processing using the Color Detection and Segmentation methodology, and the result may be seen manually by double-clicking on any section of the picture to determine the color of the diamond or automatically once the image is processed for quicker and easy sorting.

Keywords - Color detection, Diamond sorting, Machine Learning**,** OpenCV, RGB Color Model

I. INTRODUCTION

In image processing and computer vision, color is very important. It is the first stage in several computer vision applications and is essential for object identification. Color differentiation is one of the many industries that have been impacted by technological advancements. For decades, diamond mining has been a top-notch industry that continues to wow with its achievements. Even though increased technology has given it greater weight, the manual procedure of diamond sorting by people still exists. [1]. A diamond that is chemically and physically flawless is transparent, with no tint or color. However, nearly no precious stone natural diamonds are entirely flawless. Chemical contaminants or structural flaws in the crystalline lattice can change the hue of a diamond.[2]. A diamond's color may either subtract from or increase its value, guess it depends on the tint and strength of its color.

It's critical to understand the notion of color recognition before diving into the project's hypothesis. The method of finding the name of any color in a given picture is known as color detection. Emotions may be the key booster at the moment for the aforementioned color identified by humans over robots, thus people conduct this behavior naturally and without exerting any effort. However, this is not the case with computers, as computers do not display any results based on emotions, but rather on facts and particular, unlike people, who are affected by emotions and similarly colorized settings may and might elicit a feeling (colored). Light receptors in the eyes recognize the color and send the signal to the brain. This is not an exaggeration because humans have always identified certain lights with their color designations. The same strategy is effective in detecting color names in this project.

In the early 1950s, GIA created the official diamond color chart, at a period when there were many various and subjective words for defining a diamond's color, such as white, blue-white, and AAAA. The GIA spectrum starts with the letter D, which denotes colorlessness, and progresses to the letter Z, which denotes light yellow, light brown, or light grey.[3]. The GIA diamond color chart divides the 23 color categories into five sub-categories which are: (D-F) for colorless, (G-J) near-colorless, (K-M) faint, (N-R) for very light, and (L-R) (S-Z) light.[4].

The majority of detection methods rely on machine learning. Machine learning (ML) is described by Arthur Samuel as "a branch of research that provides machines with the capacity to acquire information without being explicitly programmed."

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 2309-2315 www.ijaem.net ISSN: 2395-5252

Machine Learning (ML) is the capacity of a machine to perform and improve its learning process based on its own experiences without requiring human intervention.[5]. The method begins with the provision of high-quality data, which will be used to train our machines (computers) using machine learning techniques based on the data and other ways. The algorithms we use are influenced by the sort of data we currently have and the task we want to automate. A dataset containing 865 color names and RGB values is used to train the computer how to predict a color name and its RGB value automatically using this method. The RGB color space is a color scheme in which red, green, and blue photons are blended in different ways to create a wide spectrum of colors. Each of an image's red, green, and blue color components is represented by an MxNx3 pixel matrix with M rows and N columns of pixels in the RGB model.

II. LITERATURE SURVEY

According to certain current facts that illustrate and offer tangible knowledge and comprehension of the preference of this specified project of color detection, there are some recommended strategies for the detection and recognition of colors. Duth, Sudharshan & Deepa, M. [6]. provides a way for identifying two-dimensional pictures in MATLAB via color thresholds and RGB Color model to determine the user's preferred color in the image. in their methodology, the system reads an inputted image, it then creates RGB Color bands of the Image and calculates and plots the color band histogram in red, green, and blue. After plotting the color band histogram, it subtracts the two pictures to create a 2-dimensional black and white image after converting the 3-D RGB image to a grayscale image, acknowledges the linked element mark in the linked area of the image, filters the noisy picture parts with a median filter, and computes the metric for each marking area. It's a fine strategy, but the algorithm is flawed. Extra pixels that don't appear to be part of the needed region are discovered, and these mistakes expand as the noise level rises since it just analyses the image element's intensity values and ignores any relationship between them.

In the work of N. Otsu (1979) [7] a provided method for mechanically selecting a criterion from a grey level bar graph from the perspective of discriminant analysis. This solves the problem of establishing the effectiveness of benchmarks right away. The discriminant criteria discover the optimal threshold or combination of levels by maximizing the discriminating measure character from the alphabeting addition. Wyszecki G & Stiles WS [8] present a color detection project that includes plausible elaboration and explains color science ideas and procedures. The RGB display is used to identify the shade in an image or an object. The RGB stands for Red Green Blue in a literal sense. The RGB display is a color-display paradigm that employs red, green, and blue lights to generate various effects to provide a broad or narrow spectrum of colors. The respected color science theories and policies are labeled in this planned study activity.

M. Sezgin et al [9]. In his research, he classified image threshold techniques, published their formulas in a uniform notation, and compared their performance. A histogram, spatial clusters, volatility, object properties, strong connectivity, or a surface in its native grayscale are all examples of threshold approaches. It also recognizes sophisticated threshold techniques, such as Non-Destructive Testing (NDT) and documentary image applications. Nikhil Pandey, Aayushi Saxena and Amanya Verma.[10]. Their study offered a color detection system written in Python that makes use of the OpenCV library and pandas. When a user clicks on a specific region of an image, the system recognizes the color. The color detection and image processing algorithms were effectively applied.

Neal N. Xiong, et al. [11]. In this article, they offer a novel real-time color image segmentation technique based on RGB color space color similarity. Color similarity may be calculated using the recommended color component calculation approach, which provides a color-class map, once the dominant color is determined using color and brightness information in RGB color space. The information from the corresponding color-class map is then used to classify the pixels. They also offer a color correction and light source adjustment strategy due to the possibility of measurement error. They create a novel method for detecting a fire in a live video that incorporates these characteristics by combining a recommended segmentation algorithm with a color sensor in real-time color image division for Cyber-physical systems (CPS). Experiments showed that the proposed approach for fire detection based on vision and detection in movies was effective, yielding accurate results that could be used in real-time evaluation or analysis. Masato Takahashi, et al. [12] created color charts incorporating the colors of human skin and tongue as a reference to help clinicians better recognize the color of patients during a telemedicine check-up. After a subjective review by eight medical professionals, the proposed technique was unanimously deemed practical in terms of a color

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 2309-2315 www.ijaem.net ISSN: 2395-5252

examination. The produced color chart may also be used to do automatic color correction.

Resti, Yulia & Burlian, et al (2020) [13] built and analyzed a can trash sorting system based on the CYMK (Cyan, Magenta, Yellow, and Black) digital picture color model using the k-means approach and three distinct metric distances: Manhattan, Euclidean, and Minkowski. The use of experimental results to enact three distinct metric ranges on the k-means cluster analysis system for classifying can waste into three types reveals that the average precision of k-means cluster analysis for two distances, Euclidean and Minkowski, appears to be similar, with an accuracy level difference of 1 percent or less, whereas the Manhattan distance appears to have a much lower level of accuracy. This approach is not suitable for building a can classification system since it has an average accuracy of less than 70% for both simulation and experiment data, which each incorporates three distances. P. Raguraman et al (2021).[14]. The goal of this project is to extract the color field required from an RGB image. Various procedures are implemented using the OpenCV platform. The ability to distinguish monotone colors is a major benefit of this strategy. The main purpose of this software is to recognize color hues and give an accurate prediction of their names.

Deborah T. Joy, et al (2021) [15]. According to the findings, a computer can be trained to recognize and describe colors correctly enough to be useful. The proposed color detection method uses a camera and the data it receives to determine color using RGB values. The algorithm utilized calls to a function that runs loops on realigning the distance based on the closest match. Raghav puri, et al (2018) [16]. Using Python, the Open-Source Computer Vision Library, and Numpy, the researchers developed a method for recognizing contours, forms, and colors of diverse geometrical objects in binary pictures. The photographs are processed using the most basic processes, such as loading and recognizing various shapes and colors within the given example images. Alasdair McAndrew's paper provides a detailed introduction to MATLAB-based digital image processing. [17].

III. METHODOLOGY

Color detection may be done in a variety of ways, from physical procedures to the most contemporary machine learning algorithms and even web scraping techniques. Only a few examples include the RGB model, CYMK color model, HSL color model, and other widely used color detection methods. The RGB color model is used in the proposed method for color detection. The steps are as follows:

A. Read Image

The technique begins with the input of a high-quality picture, which is then read using the OpenCV library's cv2.imread(image name) function. Any folder may be used to upload the image.

B. Import dataset.

The next step is to import the color names and RGB values from the dataset. A CSV file was utilized for this, and the CSV file was read using pandas' library.

C. RGB color extraction

This stage extracts the RGB colors from the supplied picture. An intensity value ranging from 0 to 255 is applied to each primary color. By combining three primary colors at different intensities, a variety of colors may be formed.

D. Calculate the distance of coordinated

To determine the smallest distance between all colors and compute the most matching color, the minimum distance between all colors is calculated.

E. Display the image in a window

In the rectangle window, the image with shades of color is presented. Once the double-click is initiated, the RGB values and color name are modified. To display an image, use the Cv2.imshow () function. The cv2.rectangle and cv2.putText () methods may be used to get the color name and intensity level.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 2309-2315 www.ijaem.net ISSN: 2395-5252

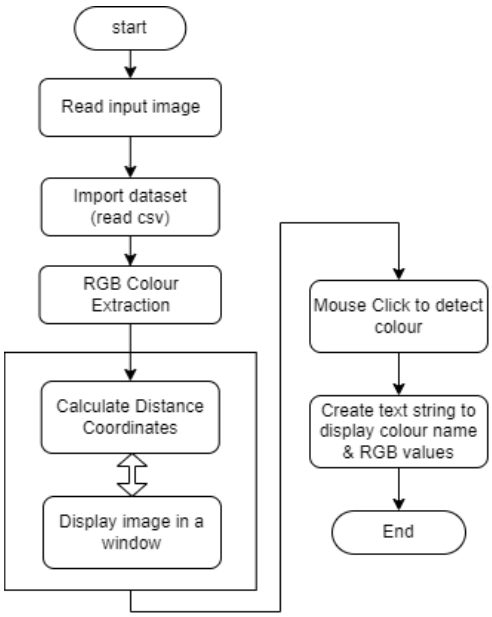

Fig. 1. System flow diagram

The graphic above depicts the flow of the detecting system as illustrated in the prior phases. The mentioned system flow improves the efficiency of the process by connecting concepts and attributes. Color identification is much easier when using RGB color combinations rather than CMYK and HSL techniques. RGB, on the other hand, has a better level of accuracy. As a consequence, the accuracy is calculated by computing the error after translating the starting RGB values to the actual color values. To make the system more user-friendly and enjoyable to use, a web application was built using the Django framework and Python as the underlying programming language, which allows a user to upload a picture, double-click it, and obtain its RGB value as well as the closest color match. The detecting method was carried out using Python modules such as OpenCV and pandas.

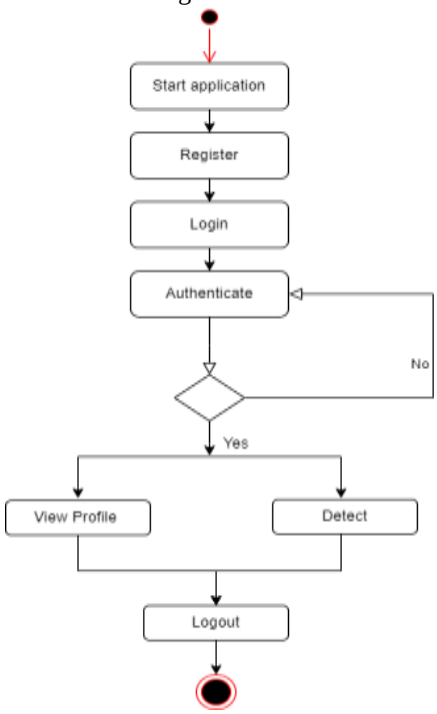

Fig. 2. Activity diagram

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 2309-2315 www.ijaem.net ISSN: 2395-5252

The activity diagram above displays the system's flow from start to finish, as well as various decision routes that occur within the activity. An activity diagram is a diagram that depicts the system's behavior. The activity diagram was used to depict both sequential and concurrent activity processing.

IV. EXPERIMENTAL RESULTS

To identify color in an image processing task, you must first obtain a picture. You may either upload photos from your hard drive or take pictures with your camera. The RGB format of an image with a resolution of MxN is a 3-D matrix of size MxNx3, with each RGB component representing a color picture in its most prominent form. The input image matrix represents the red, green, and blue portions of the image.

Fig. 3. Input images

Figure 3 shows an original photograph of diamonds from which the color must be removed. To obtain accurate information on the subject of the amount of these colors included in the picture, the cv2.imread() function in OpenCV is used to extract the primary colors from the image. The least distance between all of the colors in our dataset is calculated using the function getColorName(R, G, B) to get the most similar color.

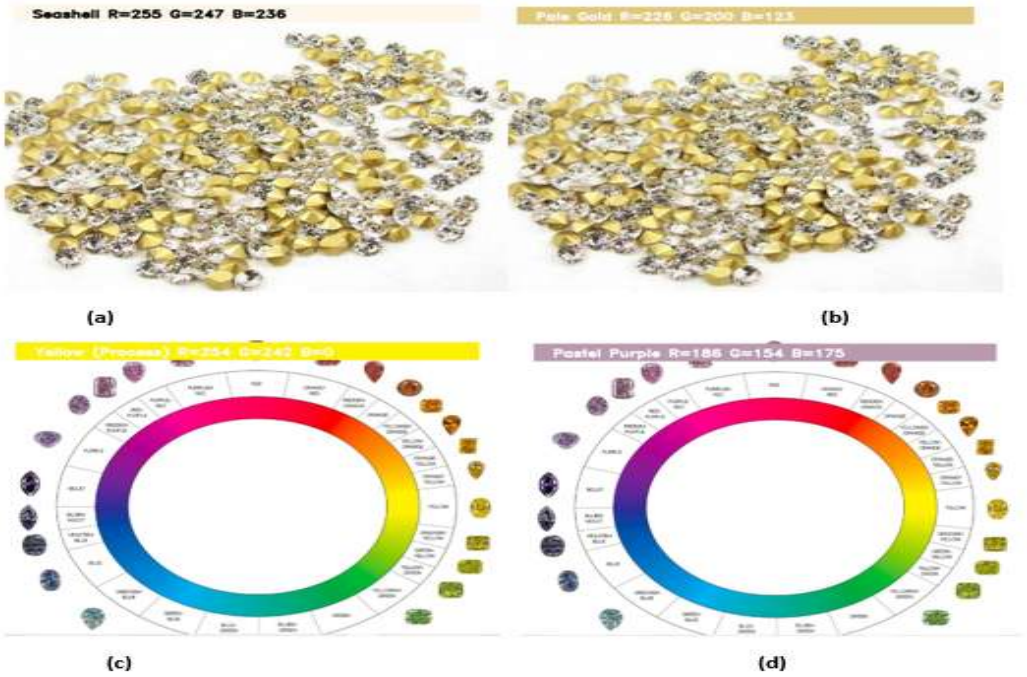

Fig. 4. The output of color detection

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 2309-2315 www.ijaem.net ISSN: 2395-5252

Figure 4 shows the color name as well as the RGB color values when you double-click on a specific area of the image. (a) Seashell RGB color intensity values (r-255, g-247, b-236), (b) Pale gold RGB color intensity values (r-226, g-200, b-123), (c) Yellow RGB color intensity values (r-254, g-242, b-0) and (d) Pastel Purple RGB color intensity values (r-186, g-154, b-175). The outcome is tested using a dataset of 865 colors and their RGB values. The final result looks to be a success, as it can predict or recognize a color with 95% accuracy. Adjusting the updated distances can assist improve the accuracy of the RGB distance calculation process when it comes to accuracy. When compared to parallel color models, the RGB model has the advantage of categorizing on all-color neutral grounds, making it easier to compute distances more often and with more accuracy.

Table 1. Dataset

Colour name	Code	R	G	B
Air Force Blue (Raf)	#5d8aa8	93	138	168
Air Force Blue (Usaf)	#00308f	0	48	143
Air Superiority Blue	#72a0c1	114	160	193
Alabama Crimson	#a32638	163	38	56
Alice Blue	#f0f8ff	240	248	255
Alizarin Crimson	#e32636	227	38	54
Alloy Orange	#c46210	196	98	16
Almond	#efdecd	239	222	205
Amaranth	#e52b50	229	43	80
Amber	#ffbf00	255	191	0
Amber (Sae/Ece)	#ff7e00	255	126	0
American Rose	#ff033e	255	3	62
Amethyst	#96c	153	102	204
Android Green	#a4c639	164	198	57
Anti-Flash White	#f2f3f4	242	243	244
Antique Brass	#cd9575	205	149	117
Antique Fuchsia	#915c83	145	92	131
Antique Ruby	#841b2d	132	27	45
Antique White	#faebd7	250	235	215
Ao (English)	#008000	0	128	0
Apple Green	#8db600	141	182	0
Apricot	#fbceb1	251	206	177
Aqua	#0ff	0	255	255

The following table depicts a portion of our dataset, the dataset used in this study is a CSV file with 865 color names, color codes, and RGB values

V. FUTURE WORK

Although the focus of this research is on color identification for diamond sorting, the system might be enhanced to increase usability in any situation. Color recognition in real-time video and the option to take a snapshot with the camera are among the features. This concept might be used in a variety of industries, including interior and exterior design. These designs are important aspects that should not be disregarded. Interior/exterior design might be tricky at times. These patterns may easily attract a lot of attention, to the point that people get intrigued about the name of a certain color. Color detection systems might potentially be used to sort mining commodities (diamonds, gold). Most paint companies utilize color-coded instructions to instruct clients on which colors to employ while developing their infrastructure. Finally, a pollution detection system might be combined with a color detecting system. Air pollution is one of the least monitored areas, and it is often discovered at inopportune times. The capacity to identify the color of the atmosphere might be immensely valuable in preventing mishaps or accidents (such as fires) from wreaking havoc on a particular area, hence reducing risk and enabling data collection. Finally, color detecting systems may be employed in both communication and information systems.

VI. CONCLUSION

We provide a method for color detection systems for diamond sorting that uses machine learning to train the system how to recognize color in RGB space and forecast its name and values in this research article. The OpenCV library is utilized for computer vision and image processing, while the Pandas library is used to acquire and analyze data from the dataset. In image processing, the Color Detection technique is accomplished using the Color Detection and Segmentation methodology. The outcome is tested using a dataset of 865 colors and their RGB values. The final result looks to be a success since it can predict or recognize a color 95% of the time.

REFERENCES

[1] Maconachie, Roy, and Tony Binns. "'Farming Miners' or 'Mining Farmers'? Diamond Mining and Rural Development in Post-Conflict Sierra Leone." Journal of Rural Studies, vol. 23, no. 3, 2007, pp. 367–380., https://doi.org/10.1016/j.jrurstud.2007.01.003.

[2] Wikipedia. "Diamond Color." Wikipedia, Wikimedia Foundation, 4 Mar. 2022, https://en.wikipedia.org/wiki/Diamond_color.
[3] GIA. "Diamond Color Chart: The Official Gia Color Scale." GIA 4Cs, 20 Oct. 2017, https://4cs.gia.edu/en-us/blog/diamond-color-chart-official-gia-color-scale/.
[4] Zhiguo Ren, et al. "Diamond Color Grading Based on Machine Vision." 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, 2009, https://doi.org/10.1109/iccvw.2009.5457523.
[5] R, Praba., et al. "Study on Machine Learning Algorithms." International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 2021, pp. 67–72., https://doi.org/10.32628/cseit2173105.
[6] Duth, Sudharshan & Deepa, M. (2018). Color detection in RGB-modeled images using MATLAB. International Journal of Engineering and Technology (UAE). 7. 29-33. 10.14419/ijet. v7i2.31.13391.
[7] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," in IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62-66, Jan. 1979, DOI: 10.1109/TSMC.1979.4310076.
[8] Wyszecki G & Stiles WS, Color science, New York: Wiley, (1982).
[9] Mehmet Sezgin, Bülent Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," J. Electron. Imag. 13(1) (1 January 2004) https://doi.org/10.1117/1.1631315.
[10] Nikhil Pandey, Aayushi Saxena, Amanya Verma. Color Detection System. International Journal of New Technology and Research (IJNTR) ISSN: 2454-4116, Volume-7, Issue-5, May 2021.
[11] Xiong, N., Shen, Y., Yang, K. et al. Color sensors and their applications based on real-time color image segmentation for cyber-physical systems. J Image Video Proc. 2018, 23 (2018). https://doi.org/10.1186/s13640-018-0258-x.
[12] Masato Takahashi1, Ryo Takahashi, Yasuhiro Morihara, Isseki Kin, Keiko Ogawa Ochiai, Norimichi Tsumura "Development of a camera based remote diagnostic system focused on color reproduction using color charts" International Society of Artificial Life and Robotics (ISAROB) (2020). https://doi.org/10.1007/s10015-020-00627-1.
[13] Resti, Yulia & Burlian, Firmansyah & Yani, Irsyadi & Rosiliani, Dinda. (2020). Analysis of a cans waste classification system based on the CMYK color model using different metric distances on the k-means method. Journal of Physics: Conference Series. 1500. 012010. 10.1088/1742-6596/1500/1/012010.
[14] P. Raguraman, A. Meghana, Y. Navya, Sk. Karishma, S. Iswarya "Color Detection of RGB Images Using Python and OpenCv" International Journal of Scientific Research in Computer Science, Engineering and Information Technology Vol-7, Issue-1, (2021) doi: https://doi.org/10.32628/CSEIT217119.
[15] Deborah T. Joy, Gitika Kaur, Aarti Chugh and Shalini B. Bajaj "Computer Vision for Color Detection" International Journal of Innovative Research in Computer Science & Technology (IJIRCST) Vol-9, Issue-3, (2021) https://doi.org/10.21276/ijircst.2021.9.3.9.
[16] Raghav puri, Archit gupta, Manas sikri "Contour, Shape, and Color detection using Open cv – Python" International Journal of Advances in Electronics and Computer Science, Vol-5, Issue-3, (2018).
[17] McAndrew A, "An introduction to digital image processing with matlab notes for scm2511 image processing", School of Computer Science and Mathematics, Victoria University of Technology, (2004).