sustainability

MDPI

Article

An Image Retrieval Framework Design Analysis Using Saliency Structure and Color Difference Histogram

Himani Chugh [1,2], Sheifali Gupta [2], Meenu Garg [2], Deepali Gupta [2], Heba G. Mohamed [3], Irene Delgado Noya [4,5], Aman Singh [4,6] and Nitin Goyal [7,*]

1 Chandigarh Group of Colleges, Landran, Mohali 140307, Punjab, India
2 Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura 140401, Punjab, India
3 Department of Electrical Engineering, College of Engineering, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
4 Higher Polytechnic School, Universidad Europea del Atlántico, C/Isabel Torres 21, 39011 Santander, Spain
5 Department of Project Management, Universidad Internacional Iberoamericana, Campeche 24560, Mexico
6 Faculty of Engineering, Universidade Internacional do Cuanza, Estrada Nacional 250, Bairro Kalua-Panda, Cuito-Bié 250, Angola
7 Computer Science and Engineering Department, Shri Vishwakarma Skill University, Palwal 121102, Haryana, India
* Correspondence: er.nitin29@ieee.org

Abstract: This paper focuses on retrieving plant leaf images based on different features that can be useful in the plant industry. Various images and their features can be used to identify the type of leaf and its disease. For this purpose, a well-organized computer-assisted plant image retrieval approach is required that can use a hybrid combination of the color and shape attributes of leaf images for plant disease identification and botanical gardening in the agriculture sector. In this research work, an innovative framework is proposed for the retrieval of leaf images that uses a hybrid combination of color and shape features to improve retrieval accuracy. For the color features, the Color Difference Histograms (CDH) descriptor is used while shape features are determined using the Saliency Structure Histogram (SSH) descriptor. To extract the various properties of leaves, Hue and Saturation Value (HSV) color space features and First Order Statistical Features (FOSF) features are computed in CDH and SSH descriptors, respectively. After that, the HSV and FOSF features of leaf images are concatenated. The concatenated features of database images are compared with the query image in terms of the Euclidean distance and a threshold value of Euclidean distance is taken for retrieval of images. The best results are obtained at the threshold value of 80% of the maximum Euclidean distance. The system's effectiveness is also evaluated with different performance metrics like precision, recall, and F-measure, and their values come out to be respectively 1.00, 0.96, and 0.97, which is better than individual feature descriptors.

Keywords: color difference histogram; saliency structure histogram; HSV; FOSF

check for updates

Citation: Chugh, H.; Gupta, S.; Garg, M.; Gupta, D.; Mohamed, H.G.; Noya, I.D.; Singh, A.; Goyal, N. An Image Retrieval Framework Design Analysis Using Saliency Structure and Color Difference Histogram. *Sustainability* **2022**, *14*, 10357. https://doi.org/10.3390/su141610357

Academic Editors: Gwanggil Jeon and Konstantinos Makantasis

Received: 10 June 2022
Accepted: 15 August 2022
Published: 19 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Plant identification is a process that consists of sorting plants into a sequence of categories according to their similarities and differences. Botanists and farmers are frequently enlisted to work with enormous collections of different plant images. They have been performing this task for thousands of years, but the process can be difficult and more time-consuming because it includes many subcategories and characteristics [1]. The recognition of plants and classification based on color images and shape features is a difficult problem for computer vision researchers because most plants do not have proper models or representations. Scientists may be able to make better use of computers to identify plants by using image processing and computer vision techniques [2]. They require automated

technologies to help them with their tasks. This study describes a system that takes an image of a leaf of a plant specimen and searches a database to find similar images [3].

In today's world, people rely on search engines such as Google, Yahoo, and Bing to find images. Provoked by market interest to explore different services, image retrieval has turned out to be a tremendously booming area [4]. In the past several years Content-Based Image Retrieval (CBIR) has been contemplated with more consideration as tremendous measures of images in different fields such as medical images, satellite images, paintings and sculpture images, business images, etc. Image data sets are typically exceptionally large and as a rule, the images are recorded and indexed by exclusive keywords given by an individual [5]. These keywords help in recovering images as per the need of the user, but generally this articulation approach is not satisfactory. This paper presents a business intelligence approach to the agricultural domain. It is inspired by two aspects of this industry: the difficulties faced by farmers during the treatment of their diseased crops when the color and shape of the leaf change due to disease and the need for plant biologists to develop automated software that extracts and analyzes the type of disease. Therefore, color and shape are the most crucial features of an image for retrieval purposes.

Considering all these factors, an image presented by a pictorial view approach makes the framework in which a comparative image to the query image is provided to the user. Many studies have proposed that CBIR makes use of the visual properties of an image for retrieval purposes. Color, texture, and shape are three characteristics by which images are retrieved. The process of image retrieval becomes more effective if the shape features are combined with other color or texture features. Most plant leaves are green or brown, but their shapes are distinct, which can be further used for the identification of leaves. So, the shape of a plant leaf is one of the most essential characteristics for distinguishing between different plant species. Therefore, the degree of closeness between an image contained in a query and a database image can be estimated by the distribution of colors, texture, or shape that appears in both images [6,7]. A lot of work has already been done using color and texture properties. Looking for images utilizing shape features has drawn much consideration. The color and shape of images are likewise a focus of the current research. To accomplish this task, CDH and SSH descriptors have been proposed. It has been shown that the combination of CDH and SSH improves the performance of the system.

The work done in this article is based on the search for plant leaves. Leaves are two-dimensional in nature and color and shape are one of the most crucial characteristics for distinguishing between different species of plants. As a result, a simple and automatic approach for correctly discriminating and recognizing the leaf forms of various species must be devised [8]. In this paper, a new feature descriptor that combines both the color and shape features is proposed for extracting the features of leaf images. Some of the major contributions are:

1. An innovative model is proposed for the retrieval of plant images in which a hybrid combination of color difference histogram (CDH) and saliency structure histogram (SSH) descriptors are used to design a detailed feature map.
2. The CDH descriptor uses the hue and saturation values of a leaf image to compute the color features.
3. The SSH descriptors use first-order statistical features (FOSF) of a leaf image to compute the shape features.
4. A new feature map is generated by concatenating features extracted from CDH and SSH descriptors to enhance the image retrieval accuracy.
5. Euclidean distance approach is used as a distance metric for measuring the similarity index between the query image and the database images.
6. The performance of the proposed model is evaluated using precision, recall, and F-measure.

The remaining paper is organized as follows: Section 2 provides a review of the relevant literature while Section 3 provides further details about our proposed model. The model is evaluated in Section 4 followed by the concluding remarks in Section 5.

2. Related Work

Color, texture, and shape are commonly employed in CBIR systems to represent any image. Color is an important aspect that is used to identify objects and retrieve images. The similarity of a leaf was measured by considering shape and venation features together. Nearest Neighbor (NN) search scheme was implemented to construct the matrix and for the computation of similarity degree. But this prototype computes the venation arrangements and similar points of intersection between different query images [8]. To retrieve images, researchers use a variety of different features. In particular, they use SURF, CHF, and BHF. These are popular features, but they are not suitable for large databases [9]. The centroid-contour distance curve and the object eccentricity (or elongation) were used in a two-step retrieval method for leaf image retrieval systems. This approach was fruitful only while considering shape features. Other local features must be adopted for better-improved results [10]. The representation of an image in terms of shape features was also implemented using the curvature scale space. The contours of an image are represented by maxima and minima but sometimes the self-intersection points did not provide the exact shape of representations [11]. Morphological image processing was implemented on plant leaves to extract and mark the veins and networks present in the plant leaves which is used to retrieve the images of plants based on their internal structure [12]. Leaf vein extraction was carried out by a two-stage procedure. First, the intensity histogram of the leaf image was used to segment the leaf into rough regions using preliminary segmentation. The second stage of this process used the trained artificial neural network (ANN) classifier, used for the detection of leaf pixels and veins extraction [13]. An improved feature, called a color correlogram, was developed that can be used to index and compare images. This feature captures the spatial correlation between different colors in an image and thus is effective in discriminating between different images [14]. Various color descriptors were used in MPEG-7 like Dominant Color Descriptor (DCD), the Color Layout Descriptor (CLD), and a scalability descriptor (SCD) for the extraction of features [15,16]. CBIR systems, which automatically extract and categorize the information from images, have used features of shapes in the images because humans are quick to recognize objects based on their shapes. A very active topic of research for retrieving similar images from vast databases based on their content is a difficult task. According to studies, finding related images based on shape is a useful technique for this. Numerous techniques are available in the literature for this aim. To identify patterns among shapes in a big dataset, a composite shape descriptor was created [17]. Researchers took various surveys based on the shape feature extraction techniques like boundary-based approaches, shape-based approaches, and Fourier-based approaches [18]. An object image retrieval method based on Compactness vectors was also presented [19]. The extraction of various features of the image was adequately acknowledged using Local Binary Patterns (LBP). An image retrieval method was proposed using directional image features and an adaptive threshold [20]. Medical images are retrieved using several different distance methods and color difference histograms [21].

From the above literature survey, it is found that a lot of work has been completed to find the similarities among various images based on color features, texture features, and shape features, individually. A variety of techniques are used based on either color mapping and indexing to find similar images based on the features of the shapes or to retrieve images based on the texture features for retrieval purposes. Very little work has been done using a combination of different features for the retrieval of images. In this study, a hybrid approach is used in which color and shape features are combined for the retrieval of images.

3. Proposed Model for Leaf Image Retrieval

The methodology proposed in this article focuses on retrieving the images from the huge collection of a dataset having different categories of leaf images. All the images present in the dataset are first loaded and are resized to the same level. The HSV and FOSF features

are extracted from each image using CDH and SSH, respectively. The features extracted from the images are combined using both descriptors. Finally, the query image is uploaded and all the features of that particular image are extracted using the same descriptor of CDH and SSH. Features computed using both the descriptors are combined using concatenation. After that, the Euclidean distance is evaluated to measure the difference between an image in the dataset and a requester image. In the end, the image retrieval operation is performed using sorting and thresholding techniques. The effectiveness of the system is computed based on different performance metrics. The flowchart of the proposed methodology is represented in Figure 1.

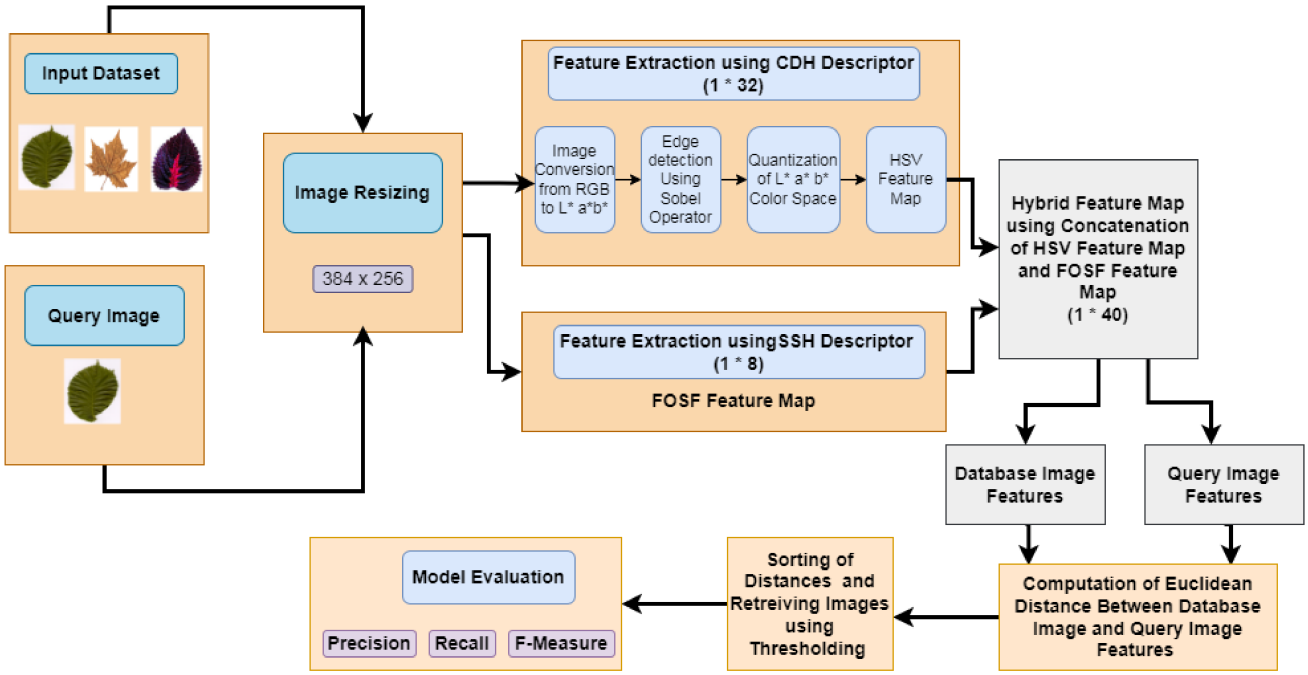


Figure 1. Proposed Model for Leaf Image Retrieval using CDH and SSH Descriptor.

3.1. Loading Dataset Images and Query Image

The experiment is performed on random images of different plant leaves with different colors and shapes. Sample images of each of the categories in the dataset are shown in Figure 2. A dataset of 100 images is created to perform the experiment. Each image belonged to one of four categories of different types of plant leaves: elm, maple, coleus, and croton. Each category contained 25 images. The images are in JPEG format each having a size of 128 by 192 pixels (or 192 by 128 pixels) [22,23]. The query image of the elm leaf is taken for simulation purposes.



(**a**) (**b**) (**c**) (**d**)

Figure 2. Categories of Leaf Images (**a**) Elm Leaf (**b**) Maple leaf (**c**) Coleus leaf (**d**) Croton.

3.2. Image Resizing

Different types of random leaf images are collected in the dataset of different colors, sizes, and shapes. All the images are required to be of the same size to create uniformity, thus image resizing is performed. This uniformity helps in providing computing accurate features. The images are resized to 384 by 256 pixels to increase the new pixel information in the image [24].

3.3. Features Extraction Using CDH Descriptor

This method employs the extraction of features of database images as well as query image using CDH descriptor as shown in Figure 3. It combines the color and edge orientation features of an image. The descriptor considers the uniform color difference and edge orientations. The algorithm to extract features from an image is described below.

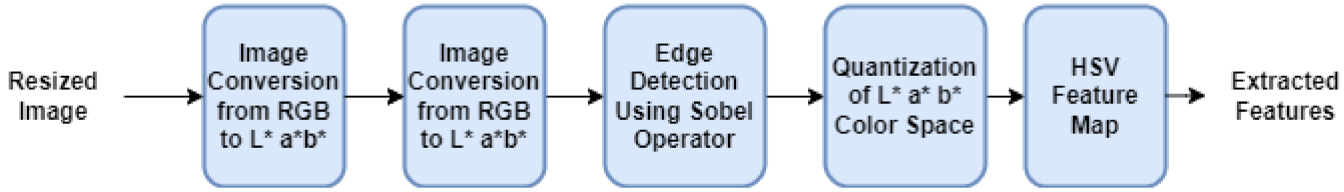


Figure 3. Method of Generating HSV Feature Map Extraction using CDH Descriptor.

3.3.1. Image Conversion from RGB to L*a*b* Color Space

The red-green-blue color model is a common color model used in graphic design and general applications. Despite its simplicity, this color space is not sufficient for mimicking human color perception. Perceptually the L*a*b* color space is uniform [25] concerning human color sensitivity. As a result, it is a better option in this situation. The RGB image is converted to L*a*b* color space as shown in Figure 4a. It is a nonlinear transformation of the RGB color space. The Lab conversion was performed using the following equations, which are derived from standard RGB to L*a*b* transformations [1,25,26]:

$$\begin{cases} ''\mathrm{L}^* = 116\left(\frac{\mathrm{y}}{\mathrm{y_n}}\right)^{1/3} - 16 & \text{for } \frac{\mathrm{y}''}{\mathrm{y_n}} > 0.00885 \\ ''\mathrm{L}^* = 903.3\left(\frac{\mathrm{y}}{\mathrm{y_n}}\right)^{1/3} & \text{for } \frac{\mathrm{y}}{\mathrm{y_n}} > 0.00885 \end{cases} \quad (1)$$

$$\mathrm{a}^* = 500\left(\mathrm{f}\left(\frac{\mathrm{x}}{\mathrm{x_n}}\right) - \mathrm{f}\left(\frac{\mathrm{y}''}{\mathrm{y_n}}\right)\right) \quad (2)$$

$$b^* = 200\left(f\left(\frac{x}{x_n}\right) - f\left(\frac{z''}{z_n}\right)\right) \quad (3)$$

with

$$\begin{cases} \mathrm{f(u)} = \mathrm{u}^{\frac{1}{3}} \text{ for } \mathrm{u} > 0.00885 \\ \mathrm{f(u)} = 7.87\mathrm{u} + \frac{16}{116} \text{ for } \mathrm{u} \leq 0.00885 \end{cases} \quad (4)$$

where

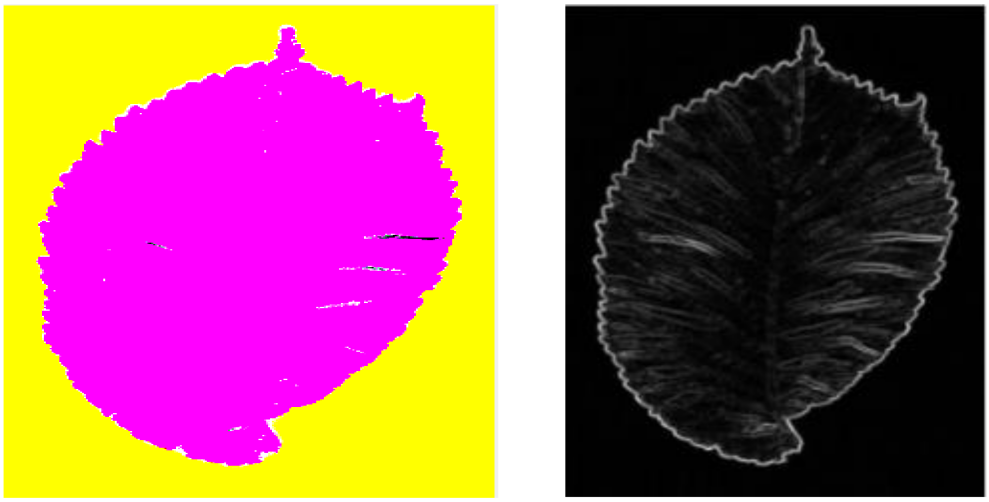
$$\begin{bmatrix} \mathrm{x} \\ \mathrm{y} \\ \mathrm{z} \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} \mathrm{R} \\ \mathrm{G} \\ \mathrm{B} \end{bmatrix} \quad (5)$$

where x_n, y_n and z_n are the values of x, y and z for illuminant (reference white point) and $\begin{bmatrix} x_n & y_n & z_n \end{bmatrix} = \begin{bmatrix} 0.950450 & 1.000000 & 1.088754 \end{bmatrix}$ [1,25].

3.3.2. Edge Detection Using Sobel Operator

The orientation of an edge is important in human perception, as it provides information on the shape and boundaries of an object. By describing an image's texture and shape, this technique also delivers semantic information. Edge orientation for each component of the L*a*b* color space is computed as presented in Figure 4b. These detection algorithms start

by converting color images to monochrome images. Then the gradient magnitude and orientation of edges in the image are determined. However, this method is less accurate than other methods such as phase-based edge detection because it loses chromatic information.



(**a**) (**b**)

Figure 4. (**a**) L*a*b* Image Converted from RGB Image (**b**) Edge Detected Image Converted from L*a*b* Image using Sobel Operator.

To preserve chromatic information, the method is applied to calculate the gradient of each color component [25]. This section uses the Sobel operators to detect horizontal (g_{xx}) and vertical (g_{yy}) edges in every pixel, then summing the computed dot product of g_{xx} and g_{yy} with the color value, creating g_{xy}. Sobel operator is preferred over other methods as it has a high-quality result on edge detection and also has the property to detect with very fast speed. Moreover, the gradient vector is composed of both magnitude and direction components. Sobel operator can be expressed using the convolution template using Equations (6) and (7).

$$g_{xx} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{6}$$

$$g_{yy} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{7}$$

The operator can be calculated by converting the edge detected image into matrix form as given by Equation (8).

$$Z = \begin{bmatrix} f(x-1,y-1) & f(x-1,y) & f(x-1,y+1) \\ f(x,y-1) & f(x,y) & f(x,y+1) \\ f(x+1,y-1) & f(x+1,y) & f(x+1,y+1) \end{bmatrix} \tag{8}$$

The horizontal direction is multiplied by the vertical direction of the template and then the vertical direction is multiplied by the horizontal direction of the template as $Fx = g_{xx}.* Z$, $Fx = g_{yy}.* Z$ provides the size of the gradient as shown by Equation (9).

$$G = \sqrt{\left({g_{xx}}^2 + {g_{yy}}^2\right)} \tag{9}$$

The direction of the gradient can be calculated using the Equation (10).

$$\varnothing = arctan\left(\frac{g_{yy}}{g_{xx}}\right) \tag{10}$$

The left side of the image is darker than the right side when the direction of the gradient is computed as zero. The output of the edge detected image using the Sobel operator is shown in Figure 4b.

3.3.3. Quantization of L*a*b* Color Space

The most important aspect of image retrieval is color, as it provides consistent spatial information. A color histogram is usually used to determine how well an image is exposed. For practical purposes, color in common computer applications is typically represented using the RGB color model, which cannot accurately portray human visual perception. To compensate for these drawbacks, the L*a*b* color space is used. In addition, each component is quantized into five bins. For instance, if n = 100, it is equal to L = 10, a = 10, and b = 10 for a color image of size m × n.

3.3.4. HSV Feature Map

Following quantization, various features are extracted from an image, including the hue, saturation, and value (HSV) histogram, autocorrelation function, and other color and wavelet moments. To calculate the color difference between two images, a color space should be used that has perceptually uniform features because it is more perceptually uniform and more analogous to the human visual system, researchers frequently utilize the HSV color space for feature extraction [21]. The remarkable homogeneity of the HSV color space makes it an especially suitable model for calculating the color difference between two points, represented by Equations (11)–(13).

$$H = \begin{cases} 60 \times \left(\frac{(G-B)}{\delta}\right). R = \max\left(R, G, B\right) \\ 60 \times \left(2 + \left(\frac{B-R}{\delta}\right)\right). G = \max\left(R, G, B\right) \\ 60 \times \left(4 + \left(\frac{R-G}{\delta}\right)\right). B = \max\left(R, G, B\right) \end{cases} \tag{11}$$

$$where\ \delta = \max\left(R, G, B\right) - \min\left(R, G, B\right)$$

$$S = \frac{\delta}{\max\left(R, G, B\right)} \tag{12}$$

$$V = \max\left(\mathrm{R}, \mathrm{G}, \mathrm{B}\right) \tag{13}$$

where H and S hold the information of chrominance and the V carries the distribution of the intensity of the given image Based on these color edge features. An edge map of R, G, and B channels separately is estimated using a histogram.

3.4. Features Extraction Using SSH Descriptor

Different features of all the images present in the database as well as the query image are extracted using (FOSF). The image is completely described at a glance by using the mean, variance, standard deviation, skewness, kurtosis, smoothness, uniformity, entropy, etc. [23]. The FOSF features of one of the random images are computed and the mathematical expressions are described in Table 1 [24].

3.5. Hybrid Feature Map Using Concatenation of HSV Feature Map and FOSF Feature Map

Different features are extracted from all the images present in the database along with the query using the CDH descriptor and the SSH descriptor for the computation of color and shape parameters. The hybrid feature map is implemented by concatenating both the descriptors. In this process, the HSV feature map extracted from the CDH descriptor is also merged with the FOSF feature map extracted from the SSH descriptor as shown in Figure 5. Overall, 32 features are computed with the help of CDH, and 8 features are computed via SSH. On applying the procedure of concatenation among both features, computed features are increased to 40, which increases the efficiency of the system for the retrieval of the query

image. Database image features and the query image features are computed to find the similarity index.

Table 1. Computed Values of FOSF Features using SSH Descriptor.

Features	Mathematical Expression	Values
Mean	$M = \bar{x} = \frac{\sum i x_i}{n}$	67.192
Variance	$V = \frac{\sum (x_i - \bar{x})^2}{n-1}$	3.56×10^3
Standard Deviation	$sd = s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$	59.696
Skewness	$SK = \frac{\sum i (x_i - \bar{x})^3}{n \times \tilde{s}^3}$ where $\tilde{s} = s \times \sqrt{(n-1)/n}$	1.107
Kurtosis	$K = \frac{\sum i (x_i - \bar{x})^4}{n \times \tilde{s}^4}$	3.536
Smoothness	$S = 1 - \left(\frac{1}{1+s^2}\right)$	0.9997
Uniformity	$U = \sum \left(\frac{\text{histogram of image}}{\text{number of elements}}\right)^2$	0.0087
Entropy	$E = -\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y) \log_2 f(x,y)$	7.326

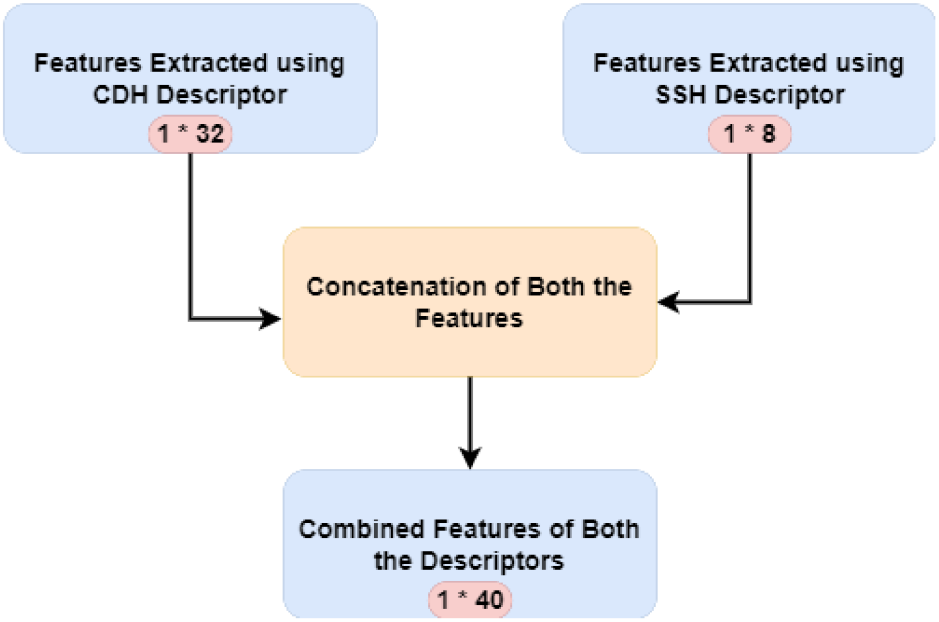


Figure 5. Combined Feature Map Extraction using Concatenation of HSV and FOSF Feature Map.

3.6. Computation of Euclidean Distance between Dataset Image and Query Image Features

Euclidean distance is the most useful method of measuring the distance between two points [27]. If there are two points (m, n) where m = (a_1, b_1) and n = (a_2, b_2), then the Euclidean distance between these points is calculated as described by Equation (14).

$$EU_d = \sqrt{(a_1 - a_2)^2 - (b_1 - b_2)^2} \quad (14)$$

3.7. Sorting of Distances and Retrieving Images Using Thresholding

The technique of image retrieval retrieves similar images to the query image by comparing the query image with every image in the database [28]. To achieve this, sorted distances between the query image and each retrieved image are considered, and then only images whose distances are less than threshold values are retained. Euclidean distance of the top 10 retrieved images is shown in Table 2. The threshold value chosen for simulation

depends upon a hit-and-trial method. A threshold is performed at different values of 40%, 75%, 80%, and 85% of the maximum distance threshold approach.

Table 2. Computed Euclidean Distance of Top 10 Images Retrieved Images in Ascending Order.

Image Index	**1**	**2**	**3**	**4**	**5**
Retrieved Image					
Computed Distance	0	2.973	4.06	4.1954	4.2329
Image Index	**6**	**7**	**8**	**9**	**10**
Retrieved Image					
Computed Distance	4.3112	4.3843	4.4677	4.7242	4.9537

4. Model Evaluation

The model is evaluated based on the implementation of different simulations. The evaluated results are explained in the following subsections.

4.1. Performance Parameters

For evaluating the performance of image retrieval systems, the model is evaluated using several measurements.

- Precision (P_r) is the ratio of the relevant images recovered to the total set of images extracted as stated by Equation (15).

$$P_r = \frac{No.\ of\ Relevant\ Images\ Extracted}{Total\ No.\ of\ Images\ Extracted} \tag{15}$$

- Recall (R_e) is computed by dividing the number of relevant images extracted by the total number of images in that dataset. The equation used to compute this parameter is represented by Equation (16).

$$R_e = \frac{No.\ of\ Relevant\ Images\ Extracted}{Total\ No.of\ Images\ in\ the\ Dataset} \tag{16}$$

- F-measure (F_m) is a combining measure of precision and recall stated by Equation (17).

$$F_m = 2 \times \frac{Pr(I_i) \times Re(I_i)}{Pr(I_i) + Re(I_i)} \tag{17}$$

4.2. Image Retrieval Results Using Only CDH Descriptor

For experiment purposes, the simulation is performed using MATLAB on the random images of different types of leaves. Sample images of elm, maple, coleus, and croton leaves are taken for simulation. The query image is selected as elm leaf on which CDH is applied at the different thresholds for the computational analysis at 40%, 75%, 80%, and 85%. Precision, recall, and F-measure is used to determine the performance of the suggested work. The comparative analysis at different thresholds based on elm leaf is represented in Table 3.

Table 3. Comparative Analysis of Elm Leaf at different Thresholds using CDH Descriptor.

Threshold Value	Retrieved Images	True Positive	Actual Images	Precision	Recall	F-Measure
40%	4	2	25	0.5	0.08	0.137931
75%	39	20	25	0.444444	0.8	0.571429
80%	45	20	25	0.512821	0.8	0.625
85%	62	19	25	0.306452	0.76	0.43678

Considering the elm leaf as the query leaf, the computed value of the F-measure is 0.625 at 80% threshold, which is observed as the best-computed value among others. It is also observed that if the threshold value is further increased to 85%, the F-measure is computed as 0.43678. Therefore 80% threshold is applied to each descriptor individually as well as in combination.

In the database of 100 images each category of the image consists of 25 images, 45 images are retrieved out of which 20 are considered to be true positive images, i.e., similar to that of the query image. At 80% threshold, precision, recall, and F-measure are computed as 40%, 80%, and 57% respectively. The computed values of all types of images at 80% threshold are shown in Table 4. The first 25 images retrieved similar to the query image after performing this experiment are shown in Figure 6.

Table 4. Performance Parameters of Image Retrieval using CDH Descriptor at 80% Threshold.

Type of Leaf Image	Retrieved Images	True Positive	Actual Images	Precision	Recall	F-Measure
Elm	45	20	25	0.444444	0.8	0.571429
Maple	38	20	25	0.526316	0.8	0.634921
Coleus	38	22	25	0.578947	0.88	0.698413
Croton	46	22	25	0.478261	0.88	0.619718

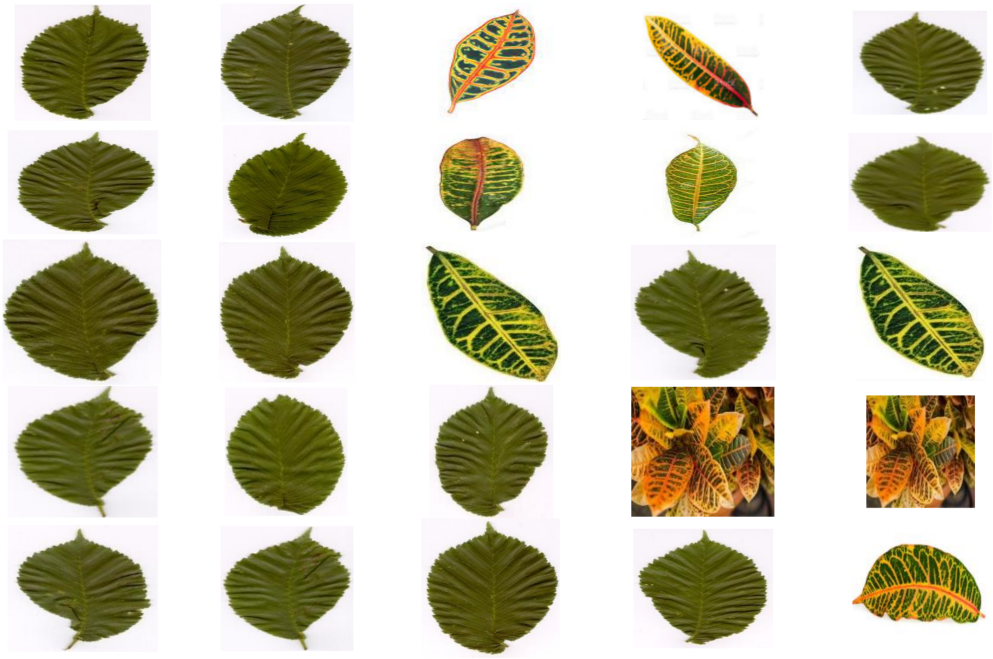


Figure 6. Retrieval Results for Top 25 Images Retrieved using CDH Descriptor.

4.3. Image Retrieval Results Using Only SSSH Descriptor

Similarly, simulation is performed on the query image of elm leaf and features are extracted using SSH descriptor. F-measure computed for elm leaf is 40.9%, which is lesser as compared to other leaves values. From the dataset of 100 images, 97 images are retrieved out of which 25 similar images of the same category are retrieved. The actual images of that

particular category are also 25. Thus, the computed value of recall is 100% and precision is 25%. Computed values of all the performance metric values like precision, recall, and F-measure for all the different categories of leaves using SSH at 80% threshold are shown in Table 5. The first 25 images retrieved similar to the query image after implementing SSH are shown in Figure 7. It can be seen when elm leaf is applied as the query image, various other categories of images are also retrieved. Based on the visual analysis also, the number of true positive images which are similar in color and shape characteristics are evaluated and computed.

Table 5. Performance Parameters of Image Retrieval using SSH Descriptor at 80% Threshold.

Type of Leaf Image	Retrieved Images	True Positive	Actual Images	Precision	Recall	F-Measure
Elm	97	25	25	0.257732	1	0.409836
Maple	97	25	25	0.257732	1	0.409836
Coleus	81	25	25	0.308642	1	0.471698
Croton	64	19	25	0.296875	0.76	0.426966



Figure 7. Retrieval Results for Top 25 Images Retrieved using SSH Descriptor.

In the database of 100 images each category of the image consists of 25 images, 45 images are retrieved out of which 20 are considered to be true positive images, i.e., similar to that of the query image. At 80% threshold, precision, recall, and F-measure are computed as 40%, 80%, and 57%, respectively. The computed values of all types of images at 80% threshold are shown in Table 6.

Table 6. Performance Parameters of Image Retrieval using Combination of CDH & SSH Descriptor at 80% Threshold.

Type of Leaf Image	Retrieved Images	True Positive	Actual Images	Precision	Recall	F-Measure
Elm	24	24	25	1	0.96	0.979592
Maple	15	12	25	0.8	0.48	0.6
Coleus	26	22	25	0.846154	0.88	0.862745
Croton	26	22	25	0.846154	0.88	0.862745

4.4. Image Retrieval Results Using Combination of CDH and SSH Descriptor

The features of the query image of the elm leaf as well as the images present in the database are extracted after combining CDH and SSH descriptors. At a threshold value of 80%, F-measure computed for the elm leaf is 0.979 as shown in Table 6, which presents the maximum performance for retrieved images similar to the query image. Out of 25 images, 24 similar images are retrieved as shown in Figure 8.



Figure 8. Retrieval Results for Top 25 Images Retrieved using Combination of CDH and SSH Descriptor.

Based on the above values computed for CDH, SSH, and the combination of CDH and SSH, the value of the F-measure is represented graphically. The performance charts for the different types of images at different thresholds are shown in Figure 9. It shows that the implementation of both CDH and SSH features combined provides better retrieval performance as compared to the individual CDH and SSH.

Based on the computed values of precision and recall the F-measure is calculated for all the different categories of leaves. When a threshold of 40% is applied, the percentage value of the F-measure comes out to be 13.7%, 48%, and 14.8% for CDH, SSH, and a combination of CDH and SSH, respectively. Similarly, for maple leaves, the computed values are 7.6%, 40.9%, and 7.6%. Considering elm leaves, the evaluated values are 68%, 48%, and 71% and for croton leaves, the F-measure are 32.2%, 34%, and 33%. It is observed that when both the descriptors are combined, it gives better results for the elm leaves as shown in Figure 9a.

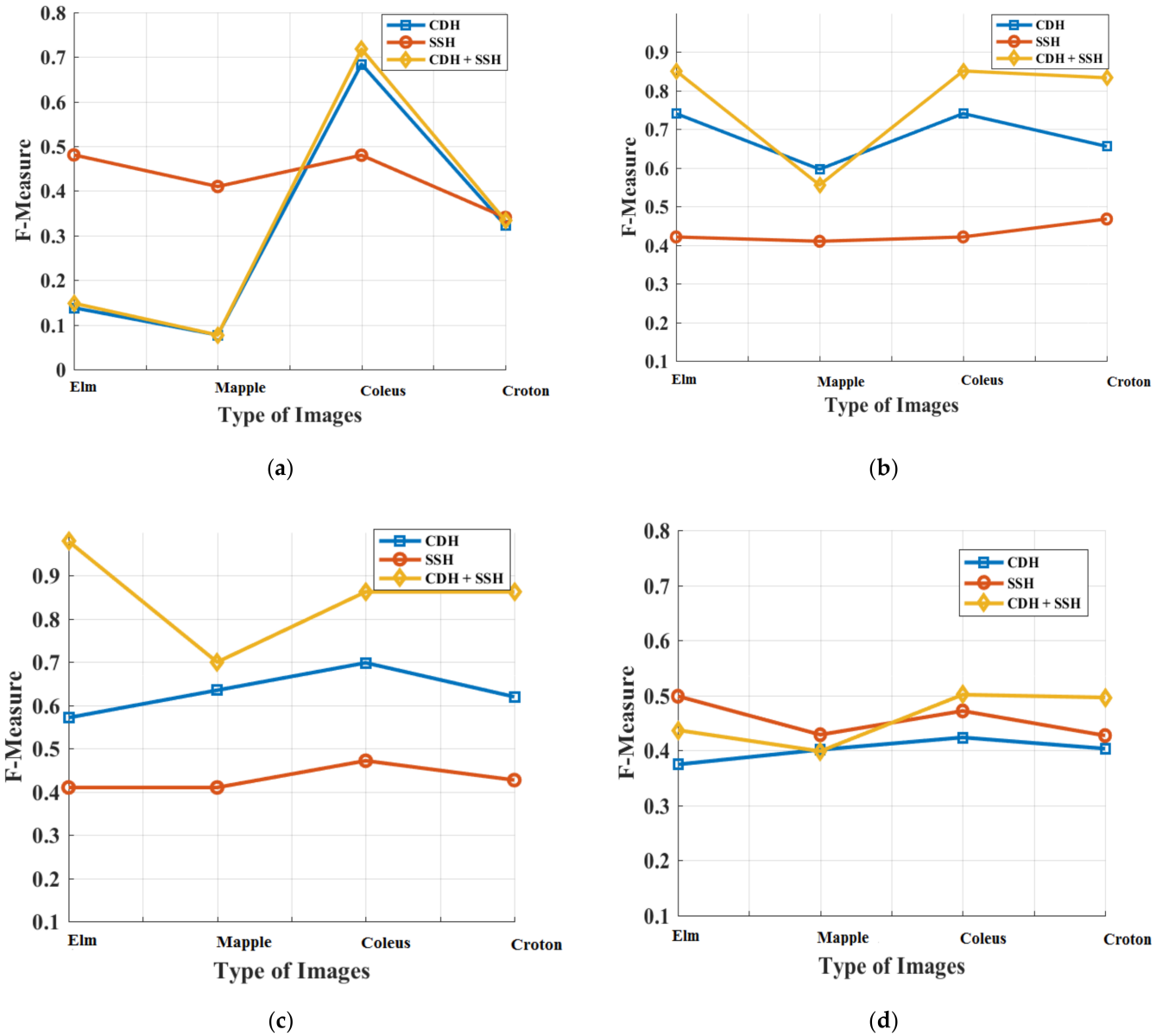


Figure 9. Performance Chart for F-Measure at (**a**) 40% Threshold (**b**) 75% Threshold (**c**) 80% Threshold (**d**) 85% Threshold.

On applying the threshold of 75%, it is observed that the values of the F-measure for elm leaf are computed as 74% using CDH, 42% using SSH, and 85% using both CDH and SSH. It is observed that after implementing both CDH and SSH the more accurate value is retrieved as shown in Figure 9b.

Further, to achieve the most appropriate results, the threshold value of 80% is applied. After applying this value, it is observed that the values computed are 69.8% using CDH, 47.1% using SSH, and 97.9% with both CDH and SSH as shown in Figure 9c for elm leaf. After applying the threshold of 85% as shown in Figure 9d, the computed values are 37.4% using CDH, 49.8% using SSH, and 43.6% using the combination of CDH and SSH. Observing the results at different threshold values, it is observed that the value of 80% is considered best for the retrieval of images.

4.5. Comparative Analysis of the Proposed Model with State-of-the-Art Techniques

The comparison of the proposed model for leaf image retrieval has been performed using CDH and SSH descriptors in terms of recall as shown in Table 7. The performance

metrics evaluated using the proposed model are compared with other state-of-the-art models. A lot of work has been done to retrieve images based on different features and patterns. The authors have used different techniques and retrieved images individually either on the shape and venation features, color features, and texture features. In the proposed model, images are retrieved using color as well as shape features. Yunyoung Nam et al. [8] achieved a recall rate of 90% with shape and venation features, Carlos et al. [28] achieved a recall rate of 81% with shape features, and B. Sathya et al. achieved a recall rate of 62% and 60% with texture features of the image. The proposed model has achieved a recall rate of 80% with color features, 100% with shape features, and 96% with the combination of color and shape features.

Table 7. Comparative Analysis of the Proposed Model with State-of-the-Art Techniques.

Sr. No.	Author(s)	Technique Used	Features Used	Recall
1	Yunyoung Nam et al. [8]	An adaptive grid-based matching algorithm	Shape & Venation features	0.90
2	Carlos et al. [28]	Contour descriptor	Shape features	0.81
3	B.Sathya Bama et al. [3]	Log-Gabor wavelet	Texture features	0.62
		Scale Invariant Feature Transform (SIFT)		0.60
4	Proposed	CDH descriptor	Color features	0.8
		SSH descriptor	Shape Features	1.00
		Combination of CDH & SSH descriptor	Color & shape features	0.96

5. Conclusions and Future Scope

In the present work, an effective image retrieval system is developed for the identification and retrieval of various leaf images. For the retrieval of the relevant leaf images according to the query image from the dataset of random leaf images, a combination of CDH and SSH descriptors is proposed here. The approach employed is demonstrated to be simple and effective and it outperforms the salient points method. The system's effectiveness is assessed at four different threshold levels, i.e., 40%, 75%, 80%, and 85% in terms of the F-measure. Based on the computed values, it is concluded that the use of hybrid feature descriptors is retrieving images accurately at a threshold of 80% which can be applied to achieve more accurate results. The proposed algorithm can also be used for the detection of relevant images in different domains like healthcare, online marketing, ancient sculptures, etc. Further work can also be more absolute if other different features of an image such as texture features will be combined along with the color and shape features of an image.

Author Contributions: Conceptualization, H.C.; methodology, S.G.; validation, M.G.; formal analysis, D.G.; investigation, N.G.; writing-original draft preparation, H.C.; writing-review and editing, D.G.; supervision, I.D.N.; project administration, H.G.M. and A.S., representation analysis, H.G.M.; All authors have read and agreed to the published version of the manuscript.

Funding: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022TR140), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Data Availability Statement: Data will be made available on request.

Acknowledgments: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022TR140), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflicts of Interest: There is no conflict of interest regarding the publication.

References

1. Patil, J.K.; Kumar, R. Plant Leaf Disease Image Retrieval Using Color Moments. *IAES Int. J. Artif. Intell. (IJ-AI)* **2013**, *2*, 36–42. [CrossRef]
2. Patil, J.; Kumar, R. Analysis of content based image retrieval for plant leaf diseases using color, shape and texture features. *Eng. Agric. Environ. Food* **2017**, *10*, 69–78. [CrossRef]
3. Bama, B.S.; Valli, S.M.; Raju, S.; Kumar, V.A. Content based leaf image retrieval (CBLIR) using shape, color and texture features. *Indian J. Comput. Sci. Eng.* **2011**, *2*, 202–211.
4. Liu, G.-H.; Yang, J.-Y. Content-based image retrieval using color difference histogram. *Pattern Recognit.* **2013**, *46*, 188–198. [CrossRef]
5. Smeulders, A.W.; Worring, M.; Santini, S.; Gupta, A.; Jain, R. Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1349–1380.
6. Ahmad, J.; Sajjad, M.; Mehmood, I.; Baik, S.W. SSH: Salient structures histogram for content based image retrieval. In Proceedings of the 2015 18th International Conference on Network-Based Information Systems, Taipei, Taiwan, 2–4 September 2015; pp. 212–217.
7. Pahalawatta, K.K. A Plant Identification System Using both Global and Local Features of Plant Leaves. Doctoral Dissertation, Department of Computer Science and Software Engineering, University of Canterbury, Christchurch, New Zealand, 2008.
8. Nam, Y.; Hwang, E.; Kim, D. A similarity-based leaf image retrieval scheme: Joining shape and venation features. *Comput. Vis. Image Underst.* **2008**, *110*, 245–259. [CrossRef]
9. Chugh, H.; Gupta, S.; Garg, M. Image Retrieval System-An Integrated Approach. *IOP Conf. Ser. Mater. Sci. Eng.* **2021**, *1022*, 012040.
10. Wang, Z.; Chi, Z.; Feng, D.; Wang, Q. Leaf image retrieval with shape features. In *International Conference on Advances in Visual Information Systems*; Springer: Berlin/Heidelberg, Germany, 2000; pp. 477–487.
11. Mokhtarian, F.; Abbasi, S. Matching Shapes With Self-Intersections: Application to Leaf Classification. *IEEE Trans. Image Process.* **2004**, *13*, 653–661. [CrossRef]
12. Soille, P. Morphological image analysis applied to crop field mapping. *Image Vis. Comput.* **2000**, *18*, 1025–1032. [CrossRef]
13. Fu, H.; Chi, Z. A two-stage approach for leaf vein extraction. In Proceedings of the International Conference on Neural Networks and Signal Processing, Istanbul, Turkey, 14–17 December 2003; Volume 14, pp. 208–211.
14. Huang, J.; Kumar, S.R.; Mitra, M.; Zhu, W.J.; Zabih, R. Image indexing using color correlograms. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, 17–19 June 1997; pp. 762–768.
15. Manjunath, B.S.; Ohm, J.R.; Vasudevan, V.V.; Yamada, A. Color and texture descriptors. *IEEE Trans. Circuits Syst. Video Technol.* **2001**, *11*, 703–715.
16. Manjunath, B.S.; Salembier, P.; Sikora, T. (Eds.) *Introduction to MPEG-7: Multimedia Content Description Interface*; John Wiley & Sons: Hoboken, NJ, USA, 2002.
17. Ahmad, J.; Jan, Z.; Khan, S.M. A fusion of labeled-grid shape descriptors with weighted ranking algorithm for shapes recognition. *arXiv* **2014**, arXiv:1406.3949.
18. Yang, M.; Kidiyo, K.; Joseph, R. A survey of shape feature extraction techniques. *Pattern Recognit.* **2008**, *15*, 43–90.
19. Achard, C.; Devars, J.; Lecassagne, L. Object image retrieval with image compactness vectors. In Proceedings of the 15th International Conference on Pattern Recognition, ICPR-2000, Barcelona, Spain, 3–7 September 2000; Volume 4, pp. 271–274.
20. Narayanan, S.S.; Vinod, D.; Athisayamani, S.; Robert Singh, A. Combination of Local Feature Extraction for Image Retrieval. In *Proceedings of Third International Conference on Sustainable Computing*; Springer: Singapore, 2022; pp. 319–328.
21. Chugh, H.; Gupta, S.; Garg, M.; Gupta, D.; Juneja, S.; Turabieh, H.; Na, Y.; Bitsue, Z.K. Image Retrieval Using Different Distance Methods and Color Difference Histogram Descriptor for Human Healthcare. *J. Health Eng.* **2022**, *2022*, 9523009. [CrossRef]
22. Ghrabat, M.J.J.; Ma, G.; Maolood, I.Y.; Alresheedi, S.S.; Abduljabbar, Z.A. An effective image retrieval based on optimized genetic algorithm utilized a novel SVM-based convolutional neural network classifier. *Hum.-Cent. Comput. Inf. Sci.* **2019**, *9*, 31. [CrossRef]
23. Li, W.; Duan, L.; Xu, D.; Tsang, I.W.-H. Text-based image retrieval using progressive multi-instance learning. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2049–2055. [CrossRef]
24. Minarno, A.E.; Suciati, N. Batik image retrieval based on color difference histogram and gray level co-occurrence matrix. *Telecommun. Comput. Electron. Control* **2014**, *12*, 597–604.
25. Gonzales, R.C.; Woods, R.E. *Digital Image Processing*, 3rd ed.; Prentice Hall: London, UK, 2007; pp. 424–447.
26. Burger, W.; Burge, M.J. *Principles of Digital Image Processing: Core Algorithms*; Springer: Berlin/Heidelberg, Germany, 2009.
27. Mehta, K.; Kumar, Y. Implementation of efficient clock synchronization using elastic timer technique in IOT. *Adv. Math. Sci. J.* **2020**, *9*, 4025–4030. [CrossRef]
28. Caballero, C.; Aranda, M.C. Plant species identification using leaf image retrieval. In Proceedings of the ACM International Conference on Image and Video Retrieval, Xi'an, China, 5–7 July 2010; pp. 327–334.