Automatic Weight of Color, Texture, and Shape Features in Content-Based Image Retrieval Using Artificial Neural Network

Akmal [a,b,*], Rinaldi Munir [a], Judhi Santoso [a]

[a] *School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia*
[b] *Department of Computer Science, Universitas Padjadjaran, Jatinangor, Indonesia*
Corresponding author: [*]akmal@unpad.ac.id

Abstract—**Image retrieval is the process of finding images in the database that are similar to the query image by measuring how close the feature values of the query image are to other images. Image retrieval is currently dominated by approaches that combine several different representations or features. The optimal weight of each feature is needed in combining the image features such as color features, texture features, and shape features. In this study, we use a multi-layer perceptron artificial neural network (MLP) method to obtain feature weights automatically and simultaneously look for optimal weights. The color moment is used to find nine color features, Gray Level Co-occurrence Matrix (GLCM) to find four texture features, and Hu Moment to find seven shape features totaling 20 neurons and all of these features become the input layer in our MLP network. Three neurons in output layers become the automatic weight of each feature. These weights are used to combine each feature's role in obtaining the relevant image. Euclidean Distance is used to measure similarity. The average precision values obtained using automatic feature weights are 93.94% for the synthetic dataset, 91.19% for the Coil-100 dataset, and 54.31% for the Wang dataset. These results have an average difference of 5.06% with the target so automatic feature weighting works well. This value is obtained at a hidden layer size of 11 and a learning rate of 0.1. In addition, the use of automatic feature weighting gives more accurate results compared to manual feature weighting.**

Keywords—**Image retrieval; color moment; hu moment; gray level co-occurrence matrix; multi-layer perceptron.**

I. INTRODUCTION

Content-based image retrieval aims to search images in huge databases based on visual image content efficiently and accurately according to user needs [1]. Image retrieval techniques have been widely researched to obtain efficient methods, but currently, there are no universally available techniques for image extraction, indexing, and retrieval. Therefore, this is still a problem that continues to be explored and, at the same time, becomes an area of research that continues to be actively carried out [2]. Hameed, Abdulhussain, and Mahmmod [3] conducted a survey, analysis, and compared the latest methodologies in the CBIR field in the 2015-2020 interval, which inspired further research in the CBIR field.

Color, texture, and shape are low-level features that are particularly popular in content-based image retrieval (CBIR) [4]. Using visual image information such as color, shape, and texture in image representation and indexing is difficult in finding and retrieving images from database collections [5]. Retrieval efficiency is still unable to meet people's needs when only one image feature is retrieved. Zenggang, Zhiwen, and Xiaowen [6] proposed a technique that combines the cumulative histogram method with shape features based on Hu moments. Pradhan et al. [7] proposed a way to overcome the ineffectiveness of visual feature-based CBIR by performing feature extraction and combining ROA and non-ROA features. This work extracts image ROA to get the image's semantic meaning based on multi-directional texture and color features based on spatial correlation [7]. Ahmed, Ummesafi, and Iqbal [8] developed a CBIR for object recognition and retrieval by fusion of spatial color features and edge features and then obtained image indexing using bag-of-visual-word and image retrieval with matching schemes. Wang et al. [9] proposed a CBIR for finding the region of the object of concern using the Itti-Koch saliency model. Next, extract the prominent image patterns using a fusion of a multi-feature scheme [9]. Multi-feature fusion was used by Dongmei Niu et al. to develop an image retrieval method that looks for color differences by extracting them using local binary patterns (LBP) after examining the direct

relationships between shape and texture features and between color and texture features [10].

In his CBIR system, Alsmadi [11] employs descriptors and features of color, shape, and texture. Kayhan and Fekri-Ershad [12] suggested a CBIR system based on a mix of weights between color features using quantization color histograms and texture features using local environmental differences patterns, modified local binary patterns, and GLCM [12]. Varish [13] proposed a feature fusion scheme that extracted shape features by calculating moments invariant of multi-resolution-based sub-images at various levels. Texture features were extracted using GLCM, and color features were extracted using the probability histogram model method. Chu and Liu [14] used a multi-integration feature model to perform CBIR by extracting color and edge data and providing a multi-integration feature histogram. Tena et al. [15] conducted a study on the application of feature extraction methods for fabric images, which were divided between conventional techniques and convolutional neural networks. Khan et al. [16] suggested a CBIR approach for image retrieval in a multi-class situation based on a hybrid features descriptor using the genetic algorithm (GA) and SVM classifier. Garg and Dhiman [17]perform the CBIR process by extracting and reducing several features to obtain a multilevel decomposition of the image using the GLCM feature and the LBP variant. Subramanian et al. [18] proposed a way to increase the efficiency and effectiveness of the CBIR system by combining color, greyscale, texture, shape, and random forest classifier features using the Particle Swarm Optimization (PSO) algorithm to select informative features. Jardim et al. [19] developed an image retrieval system for the image property industry by implementing multi-phase deep learning and image processing techniques through image signatures to provide an image representation that is close to precise so that plagiarism can be found from trademarks owned by the industry.

CBIR systems are often composed of low-level features that are constructed by giving all aspects fixed or equal weights. However, based on the image's content, it extracts information about various aspects regardless of their weightage.

In many CBIR systems, the user must manually determine the relevance of low-level characteristics. The weights supplied by the users determine the quality of performance. However, determining the right weights for the attributes is quite difficult if the user is inexperienced. There are two techniques for assigning weights in CBIR: a linear method that describes a linear combination of each feature coefficient in many iterations and a nonlinear method that combines various aspects of the AI method and the Relevance Feedback (RF) method [20].

In this paper, to manually replace the input given by the user, the weighting is automated using the artificial neural network (ANN) method. Modeling an Artificial neural network is based on a human neural network, a network of a group of processing units. ANN can change its structure based on external and internal information to solve problems. The input, process, and output layers comprise the three layers (layers) of ANN. The input layer contains variables for input data, the process layer contains object recognition steps, and the output layer contains object introduction results.

The paper is organized into four sections. The first section is this introduction, containing background and reviews of some related works. The second section will explain the proposed methods to perform automatic weight in image retrieval, including low-level feature methods and artificial neural networks. The third section will present the experimental results and discussion. Finally, the final section will discuss the conclusion and future work.

II. MATERIAL AND METHOD

The proposed method uses a combination of several methods, namely Color Moment, GLCM, and Hu Moment, which are methods for extracting color, texture, and shape features, respectively. In this case, the weighting of the combination of features is done automatically using ANN, and the results will also be compared with the weighting that is done manually.

When a user enters a query image, the system will extract color, texture, and shape features, the methods for which will be discussed in the following section. The feature extraction results are saved in a feature vector normalized, and the weight of each feature is automatically searched using ANN. Euclidean distance is used to calculate similarity, and a result image relevant to the given query image is generated, as shown in Fig 1.

A. Color Moments

The Visual features (color, shape, and texture) and semantic features are two things that are often used in the CBIR method. The visual color feature is one of the most crucial main features [21]. The extraction features of the color are easier compared to other features. In addition, color features do not depend on image transformations such as rotation, scaling, vision, and other deformations, and they have strong resilience.
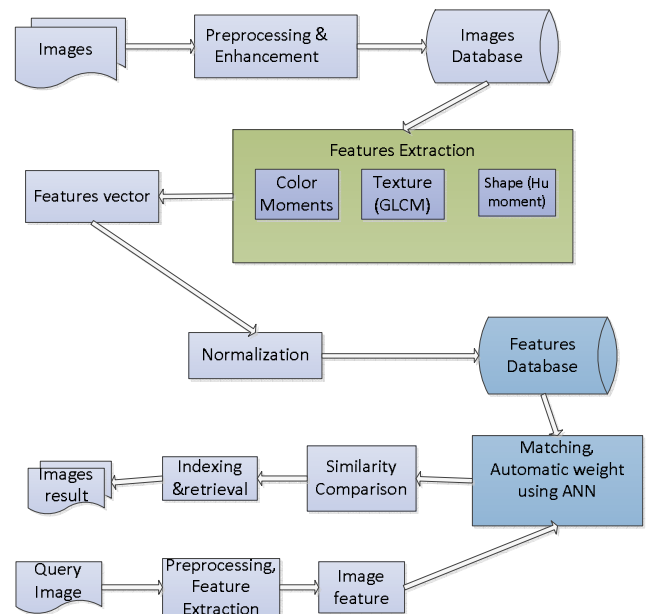

Fig. 1 Proposed content-based image retrieval approach.

The mean, variance, and skewness are the first, second, and third-order color moments that effectively and efficiently

represent the distribution of the image color. In image retrieval systems, the color moment is commonly utilized [22]

The following are the three color moments:

a. Mean color moment: the average value of the image color with p_{ij}. is the i-th color channel at the j-th image pixel.

$$E_i = \sum_{j=1}^{N} \frac{1}{N} p_{ij} \qquad (1)$$

b. Standard Deviation moment: the variance distribution's square root

$$\sigma_i = \sqrt{\left(\frac{1}{N} \sum_{j=1}^{N} (p_{ij} - E_i)^2\right)} \qquad (2)$$

c. Skewness moment: a measure of the distribution's degree of asymmetry.

$$s_i = \sqrt[3]{\left(\frac{1}{N} \sum_{j=1}^{N} (p_{ij} - E_i)^3\right)} \qquad (3)$$

B. Hu Moments

One of the fundamental characteristics of depicting objects is their shape. The use of shape features will increase the accuracy and efficiency of the image retrieval process [23]. Shape characteristics, such as an approximation of objects by a collection of shapes, have been employed in numerous CBIR systems [24]. Hu Moments is a form feature search method that consists of seven values calculated using a central moment whose values are not affected by image transformation [25]. The first six moment values are unchanged concerning translation, rotation, scale, and reflection. The seventh-moment changes to the reflection of the image.

The Hu moment is calculated using the following formula:

$$m_0 = \eta_{20} + \eta_{02} \qquad (4)$$
$$m_1 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \qquad (5)$$
$$m_2 = (\eta_{30} + 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \qquad (6)$$
$$m_3 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \qquad (7)$$

$$
\begin{aligned}
m_4 = {} & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\
& - 3(\eta_{21} + \eta_{03})^2] \\
& + (3\eta_{21} - \eta_{03})[3(\eta_{30} + \eta_{12})^2 \\
& - (\eta_{21} + \eta_{03})^2]
\end{aligned} \qquad (8)
$$

$$
\begin{aligned}
m_5 = {} & (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \\
& + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})]
\end{aligned} \qquad (9)
$$

$$
\begin{aligned}
m_6 = {} & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\
& - 3(\eta_{21} + \eta_{03})^2] \\
& + (\eta_{30} - 3\eta_{12})(\eta_{21} \\
& + \eta_{03})[3(\eta_{30} + \eta_{12})^2 \\
& - (\eta_{21} + \eta_{03})^2]
\end{aligned} \qquad (10)
$$

where

$$\eta_{ij} = \frac{\mu_{ij}}{\mu_{00}^{(i+j)/2+1}} \qquad (11)$$

Which is a normalization of the following central moment:

$$\mu_{ij} = \sum_x \sum_y (x - \bar{x})^i (y - \bar{y})^j I(x,y) \qquad (12)$$

$\bar{x}$ and $\bar{y}$ are the centroid coordinates for an area.

C. Grey Level Co-occurrence Matrix

A texture feature is a type of feature that is not based on brightness or color. This feature contains surface information and the environment around the image, and the image's spatial information can be depicted quantitatively [26]. The GLCM method is a statistical method for finding image texture features that are calculated using a gray degree distribution. The GLCM approach calculates an image region's contrast, roughness, and granularity based on the adjacent connection between pixels. Image histograms are used for first-order feature extraction, while a co-occurrence matrix is used for second-order feature extraction. The co-occurrence matrix represents the various orientation directions and spatial distances [27]. Texture analysis with GLCM has fourteen parameters [28]. Huang et al. [29] found that only four unrelated image features could be used to extract the texture features: energy, correlation, contrast, and entropy.

a. The second angular moment (energy) shows the uniformity and roughness of the image distribution.

$$f_{ASM} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_d^2 (i,j) \qquad (13)$$

b. The contrast shows image clarity and texture depth.

$$f_{CON} = \sum_{n=0}^{L-1} n^2 \left\{ \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_d (x,y) \right\} \qquad (14)$$

c. The correlation shows the degree of similarity between components in the row or column direction.

$$f_{COR} = \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} ij p_d(i,j) - \mu_1 \mu_2}{\sigma_1^2 \sigma_2^2} \qquad (15)$$

d. Entropy is used to show the randomness of the image texture

$$f_{ENT} = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_d(i,j) \log (p_d(i,j)) \qquad (16)$$

Texture features with rotational invariance are created by using offset parameters in four directions (0◦, 45◦, 90◦, 135◦). Each characteristic index is searched, and its mean and variance are calculated, resulting in a textural feature vector with no relationship to the direction.

D. Multi-Layer Perceptron Neural Network

Artificial neural network engineering has numerous applications in fields such as medicine, science, computing, engineering, nanotechnology, mining, agriculture, environment, climate, business, art, etc. Abiodun et al. [30] explained that when applied to human problems, neural network models such as feed-forward and feedback propagation artificial neural networks perform better.

A common type of ANN model is feedforward neural networks (FFNN), which can view and estimate computational models using parallel layered structures [31]. This structure consists of a series of fully connected layers and a set of neurons as processing elements. Multi-Layer

Perceptron (MLP) is a subset of FFNN consisting of three parallel layers: input, hidden, and output.
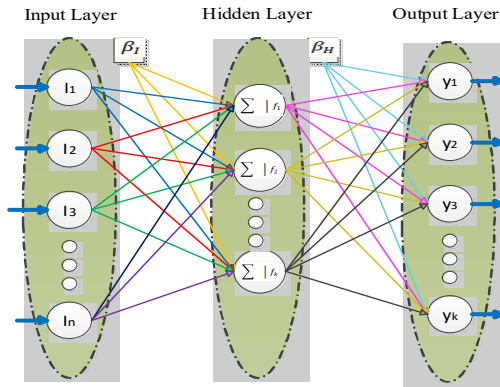

Fig. 2 MLP Neural Network

Fig 2 shows an MLP neural network with one hidden layer. Several weights in the range [1, 1] should be used to define the joints between layers of MLP. Each node performs two tasks: summarization and activation. The product of the input, weight, and bias are added in Eq (17).

$$S_j = \sum_{i=1}^{n} \omega_{ij} I_i + \beta_j \qquad (17)$$

where n represents the number of inputs, I_i represents the input variable, β_j represents the bias rate, and ω_{ij} represents the connection weight. Second, start the activation function (18) using the output of Eq (17). The S-shaped curved sigmoid function has been the most widely used type in MLP [32]. Eq (18) describes this function.

$$f_j(x) = \frac{1}{1 + e^{-s_j}} \qquad (18)$$

As a result, Eq. (19) can be used to calculate the final output of neuron j:

$$y_i = f_j\left(\sum_{i=1}^{n} \omega_{ij} I_i + \beta_j\right) \qquad (19)$$

The input layer receives features, including color, hu moment, and glcm features. The hidden layer is obtained from a trial-and-error process, and there are three neurons in output layers which will be the weight values of the fusion of the image retrieval model.

The learning steps are carried out to improve and update the network weights based on the ANN structure design. These weights are rationalized to estimate results and minimize errors.

III. RESULT AND DISCUSSION

This section will discuss the suggested system's findings. The suggested system's implementation is written in Python, and the evaluation metric is used to assess the system's performance. Precision is a popular metric for assessing a proposed image retrieval system in which the relevant image serves as the foundation for computations. Precision is defined as the percentage of relevant responses provided by the system among the retrieved occurrences.

$$P = \frac{TP}{TP + FP} \qquad (20)$$

P stands for Precision. The relevant images (True Positive) recovered are denoted by TP, while the total number of images retrieved (True Positive + False Positive) is denoted by TP + FP.

$$AP = \sum_{i=1}^{n} \frac{P_i}{n} \qquad (21)$$

AP stands for Average Precision. The mean accuracy of the class's i images (P_i), where n is the total number of images, is used to calculate AP.

$$MAP = \sum_{j=1}^{m} \frac{AP_j}{m} \qquad (22)$$

MAP represents the Mean Average Precision, AP_j represents the average precision of the j class image, and m represents the total number of classes in the database.

A. Dataset

The image database in this experiment is separated into three categories of datasets: synthetic, COIL-100, and wang (Table 1). The synthetic dataset has 200 images in four different classes with 64 x 64-pixel sizes. Dataset COIL-100 contains 100 objects with 7,200 images on a black background. Each object is placed on the table and rotated at 5-degree intervals, resulting in 72 images for each object. Wang's dataset contains 1,000 images divided into ten classes containing 100 images, 256x386 and 384x256 pixels in JPEG format consisting of classes Africans, beaches, monuments, buses, dinosaurs, elephants, flowers, horses, mountains, and food.

TABLE I
EXAMPLE OF THE THREE TYPES OF DATASET

Data set	Images Example
synthetic	
Coil-100 (Columbia Object Image Library)	
Wang	

B. Image Retrieval Prototype Model Using ANN

Giving weights $\omega_1, \omega_2,$ and ω_3 to the color moment, Hu Moment, and GLCM methods will give the combined similarity value. In this scenario, all feature vector values have been normalized so that the feature values are in the range [0..1].

$$sim = \omega_1 d_{mom}(A,B) + \omega_2 d_{hu}(A,B) + \omega_3 d_{glcm}(A, B) \tag{23}$$

Distance between two feature images measured using Euclidean Distance for both color moment, hu Moment, and GLCM, namely

$$d_x(A,B) = \sqrt{\sum_{i=0}^{n}(F_i^B - F_i^A)^2} \tag{24}$$

Weights $\omega_1, \omega_2, \omega_3$ in equation (23) can be obtained automatically using the multi-layer perceptron neural network method. Details of the design of the model as shown in Table 2.

TABLE II
DESIGN OF PREDICTED MULTILAYER ANN MODEL

Parameter	Quantity	Information
Input Layer	20 neuron	9 color moments 7 Hu moment 4 GLCM
Hidden Layer	Trial and error	8-13
Output Layer	3 neuron	Weight of color, shape, and texture
Initial weight	Trial and error	Random number
Learning Rate	Trial and error	0.1 - 0.3
epoch	Trial and error	500 - 2000
Activation Function	1	logsig

The input layer receives 20 features, which include 9 color features, 7 hu moment features, and 4 glcm features. The hidden layer is obtained from a trial and error process with the number of neurons being tested between 8 to 13 neurons, meanwhile, there are three output layers which will be the weight values $\omega_1, \omega_2, \omega_3$.

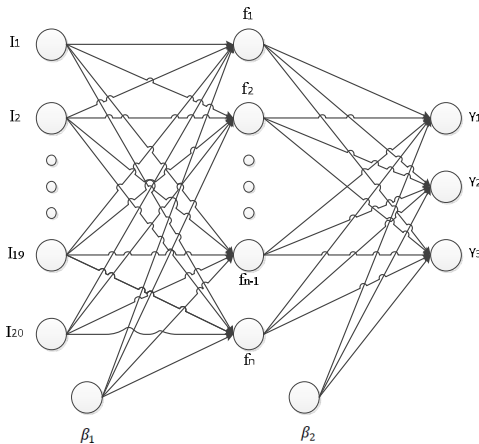From table 2, the architecture of the ANN Model can be described as shown in Fig 3.

Fig. 3 Architecture of ANN Model.

The algorithm to get the automatic weights for image retrieval proposed is as follows (Fig.4).

```
1. Obtain composite feature vectors for moment color,
   texture, and shape image query.
   Look for ColorMomentFeature
   Look for HuMomentFeature
   Look for GLCMFeature

2. Repeat step 1 for all images to obtain features in
   the database.

3. Calculate the similarity matrix using the Euclidean
   distance.
   Calculate and normalize color moment similarity
   Calculate and normalize HuMoment similarity
   Calculate and normalize GLCM similarity

4. Use ANN to get weight values (ω₁, ω₂, ω₃) for the
   color moment, hu Moment, and GLCM features
   clf = MLPClassifier(hidden_layer_sizes=11,
       learning_rate_init=0.1, max_iter=2000)
   clf.fit(x_train, y_train)

5. Color moment, Hu Moment, and GLCM features are
   combined into a single vector for similarity
   measurements.
   Similarity_total ← ω₁* similarity_moment + ω₂*
       similarity_ hu Moment + ω₃*
       similarity_glcm

6. Take the image based on the similarity index.
```

Fig. 4 Algorithm of image retrieval using automatic weight

C. Result Analysis

Table 3,4, and 5 compares MAP using automatic weighting and manual weighting to determine the best target MAP value that can be obtained. The MAP Best Target value is the value obtained by selecting the best of 67 alternative weight parameters $\omega_1, \omega_2, \omega_3$ given manually. The weight of $\omega_1, \omega_2, \omega_3$ selected is a combination of fractional values that total 100%, namely pairs (0,0,1), (0,0.1,0.9), (0,0.2,0.8), (0,0.3,0.7), …, (0.8,0,0.2), (0.8,0.1,0.1), (0.8,0.2,0), (0.9,0,0.1), (0.9,0.1,0), (1,0,0). The number of retrieved data consists of n = 10, 20, and 30 images. For the MLP parameter, the value of hidden layer size = 11 and learning rate = 0.1 is chosen, which gives the best MAP value.

TABLE III
AVERAGE PRECISION (%) OF SYNTHETIC DATASET

Methods	n=10	n=20	n=30	Average
Best Target	**97.9**	**95.7**	**91.88**	**95.16**
CM	76.55	65.8	59.71	67.35
GLCM	89.65	82.72	76.5	82.96
HM	71.5	59.1	50.46	60.35
CM+GLCM	91.85	87.4	83	87.42
CM+HM	82	71.19	63.98	72.39
GLCM+HM	90.35	81.9	75.76	82.67
CM+GLCM+HM	92.65	87.17	81.91	87.24
Automatic	**95.1**	**95.49**	**91.23**	**93.94**

CM=Color Moment.GLCM=Grey Level Co-occurrence Matrix,HM=Hu Moment,

Methods	n=10	n=20	n=30	Average
Best Target	**99.3**	**96.49**	**94.19**	**96.66**
CM	96.66	93.33	90.92	93.64
GLCM	77.79	70.99	66.38	71.72
HM	67.08	59.85	54.41	60.45
CM+GLCM	94.58	91.28	87.81	91.22
CM+HM	97.04	92.53	89.79	93.12
GLCM+HM	81.94	75.81	71.49	76.41
CM+GLCM+HM	95.45	91.62	88.42	91.83
Automatic	**92.79**	**90.95**	**89.84**	**91.19**

CM=Color Moment. GLCM=Grey Level Co-occurrence Matrix, HM=Hu Moment,

TABLE V
AVERAGE PRECISION (%) OF WANG DATASET

Methods	n=10	n=20	n=30	Average
Best Target	**61.37**	**65.69**	**61.37**	**62.81**
CM	55.72	49.89	46.99	50.87
GLCM	43.35	39.11	37.05	39.84
HM	28.46	23.71	21.99	24.72
CM+GLCM	54.67	51.25	48.82	51.58
CM+HM	55.99	50	46.97	50.99
GLCM+HM	43.68	39.08	37.12	39.96
CM+GLCM+HM	60.32	54.69	51.32	55.44
Automatic	**57.65**	**53.53**	**51.75**	**54.31**

CM=Color Moment. GLCM=Grey Level Co-occurrence Matrix, HM=Hu Moment,

Tables 3, 4, and 5 show that automating feature weights can improve image retrieval precision over manually assigning weights. The total of the combined weights is 1. As a result, weights for a single CM, HM, or GLCM are 1, 0.5, and 0.5 for a combination of two methods, and 0.33, 0.33, and 0.34 for a combination of three methods. Fig 5 depicts a graph of the MAP value of each method. It can be seen that the fusion of 2 or 3 methods will give a better precision value than using only one single method.
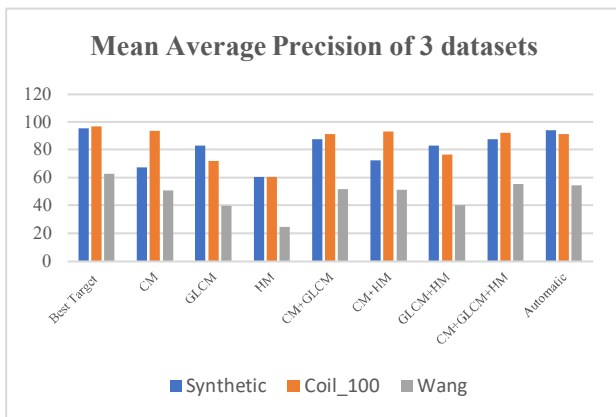

Fig. 5 Mean Average Precision (%) of 3 datasets

Furthermore, the image retrieval results will be shown based on the query image provided in Fig 6, Fig 7, and Fig 8. The image is considered correct if it belongs to the same category as the query. A predetermined (n) number of images are retrieved in ascending order for each query based on the Euclidean distance between the query and the captured images. The top ten images were retrieved using three sample image queries from 3 datasets by applying the proposed image retrieval method with both automatic and manual weighing. Generally, the proposed method with automatic weighting gives more precise results than the single method, a combination of 2 methods, or a combination of 3 methods whose total weight is 1.
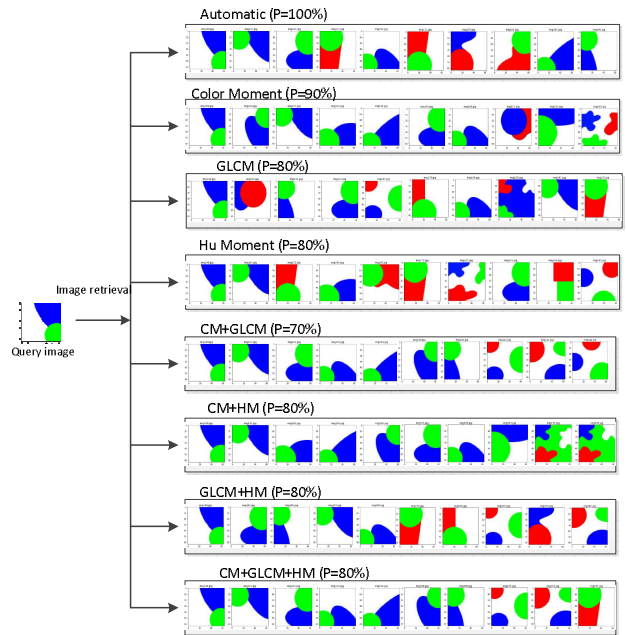

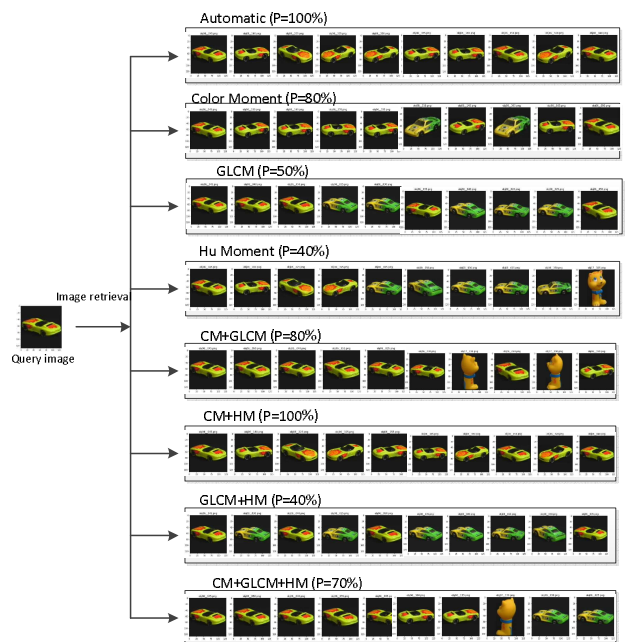Fig. 6 Example of the image retrieval results of the synthetics data set

Fig. 7 Example of the image retrieval results of the COIL-100 data set

Fig. 8 Example of the image retrieval results of theWang data set

IV. CONCLUSIONS

The content-based image retrieval system is proposed using a combination of color, texture, and shape features with automatic weighting using ANN. The results were compared with manual weighing. The feature extraction method used is the color moment, GLCM, and Hu Moment. The initial experimental results show that the fusion of 2 or 3 methods will give a better precision value than using only one single method. Automating feature weights can increase the precision of image retrieval compared to manually assigning weights.To improve retrieval accuracy, future work will involve experimenting with other various methods of color features, texture features, and other shape features.

ACKNOWLEDGMENT

REFERENCES

[1] A. Latif *et al.*, "Content-based image retrieval and feature extraction: A comprehensive review," *Math. Probl. Eng.*, vol. 2019, 2019, doi: 10.1155/2019/9658350.

[2] S. Chavda and G. Mahesh, "Content-Based Image Retrieval: The State of the Art," *Int. J. Next-Generation Comput.*, vol. 10, no. December, pp. 193–212, 2019.

[3] I. M. Hameed, S. H. Abdulhussain, and B. M. Mahmmod, "Content-based image retrieval: A review of recent trends," *Cogent Eng.*, vol. 8, no. 1, 2021, doi: 10.1080/23311916.2021.1927469.

[4] H. Ji-Zhao, G.-H. Liu, and S.-X. Song, "Content-based image retrieval using color volume histograms," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 33, no. 9, p. Article ID 6283987, 2019.

[5] B. Patel, K. Yadav, and D. Ghosh, "Current Trend and Methodologies of Content-Based Image Retrieval: Survey," 2021, doi: https://doi.org/10.1007/978-981-15-6707-0_64.

[6] X. Zenggang, T. Zhiwen, and C. Xiaowen, "Research on Image Retrieval Algorithm Based on Combination of Color and Shape Features," *J. Signal Process. Syst.*, vol. 93, no. 2, pp. 139-146., 2019.

[7] J. Pradhan, A. K. Pal, H. Banka, and P. Dansena, "Fusion of region based extracted features for instance-and class-based CBIR applications," *Appl. Soft Comput.*, vol. 102, p. 107063, 2021.

[8] K. T. Ahmed, S. Ummesafi, and A. Iqbal, "Content based image retrieval using image features information fusion," *Inf. Fusion*, vol. 51, pp. 76–99, 2019.

[9] H. Wang, Z. Li, Y. Li, B. Gupta, and C. Choi, "Visual saliency guided complex image retrieval," *Pattern Recognit. Lett. Elsevier*, vol. 130, pp. 64–72, 2020.

[10] D. Niu, X. Zhao, X. Lin, and C. Zhang, "A novel image retrieval method based on multi-features fusion," *Signal Process. Image Commun.*, vol. 87, p. 115911, 2020, doi: 10.1016/j.image.2020.115911.

[11] M. K. Alsmadi, "Content-Based Image Retrieval Using Color, Shape and Texture Descriptors and Features," *Arab. J. Sci. Eng.*, vol. 45, no. 4, pp. 3317–3330, 2020, doi: 10.1007/s13369-020-04384-y.

[12] N. Kayhan and S. Fekri-Ershad, "Content based image retrieval based on weighted fusion of texture and color features derived from modified local binary patterns and local neighborhood difference patterns," *Multimed. Tools Appl.*, vol. 80, no. 21, pp. 32763–32790, 2021.

[13] N. Varish, *A modified similarity measurement for image retrieval scheme using fusion of color, texture and shape moments*, vol. 81, no. 15. Multimedia Tools and Applications, 2022.

[14] K. Chu and G. H. Liu, "Image Retrieval Based on a Multi-Integration Features Model," *Math. Probl. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/1461459.

[15] S. Tena, R. Hartanto, and I. Ardiyanto, "Content-based image retrieval for fabric images: A survey," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 23, no. 3, pp. 1861–1872, 2021, doi: 10.11591/ijeecs.v23.i3.pp1861-1872.

[16] U. A. Khan, A. Javed, and R. Ashraf, "An effective hybrid framework for content based image retrieval (CBIR)," *Multimed. Tools Appl.*, vol. 80, no. 17, pp. 26911–26937, 2021, doi: 10.1007/s11042-021-10530-x.

[17] M. Garg and G. Dhiman, "A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants," *Neural Comput. Appl.*, vol. 33, no. 4, pp. 1311–1328, 2021.

[18] M. Subramanian, V. Lingamuthu, C. Venkatesan, and S. Perumal, "Content-Based Image Retrieval Using Colour, Gray, Advanced Texture, Shape Features, and Random Forest Classifier with Optimized Particle Swarm Optimization," *Int. J. Biomed. Imaging*, vol. 2022, 2022, doi: 10.1155/2022/3211793.

[19] S. Jardim, J. António, C. Mora, and A. Almeida, "A Novel Trademark Image Retrieval System Based on Multi-Feature Extraction and Deep Networks," *J. Imaging*, vol. 8, no. 9, 2022, doi: 10.3390/jimaging8090238.

[20] N. K. Rout, M. Atulkar, and M. K. Ahirwal, "A review on content-based image retrieval system: Present trends and future challenges," *Int. J. Comput. Vis. Robot.*, vol. 11, no. 5, pp. 461–485, 2021, doi: 10.1504/IJCVR.2021.117578.

[21] S. Lu and B. Wang, "An image retrieval algorithm based on improved color histogram," 2019, doi: 10.1088/1742-6596/1176/2/022039.

[22] R. M. Bommisetty, O. Prakash, and A. Khare, *Keyframe extraction using Pearson correlation coefficient and color moments*, vol. 26, no. 3. Springer Berlin Heidelberg, 2020.

[23] G. Xie, B. Guo, Z. Huang, Y. Zheng, and Y. Yan, "Combination of Dominant Color Descriptor and Hu Moments in Consistent Zone for Content Based Image Retrieval," *IEEE Access*, vol. 8, pp. 146284–146299, 2020, doi: 10.1109/ACCESS.2020.3015285.

[24] S. Deniziak and T. Michno, "New content based image retrieval database structure using query by approximate shapes," *Proc. 2017 Fed. Conf. Comput. Sci. Inf. Syst. FedCSIS 2017*, vol. 11, pp. 613–621, 2017, doi: 10.15439/2017F457.

[25] Z. Huang and J. Leng, "Analysis of Hu's moment invariants on image scaling and rotation," *ICCET 2010 - 2010 Int. Conf. Comput. Eng. Technol. Proc.*, vol. 7, no. May 2010, 2010, doi: 10.1109/ICCET.2010.5485542.

[26] A. Humeau-Heurtier, "Texture feature extraction methods: A survey," *IEEE Access*, vol. 7, no. c, pp. 8975–9000, 2019, doi: 10.1109/ACCESS.2018.2890743.

[27] A. Eleyan and H. Demirel, "Co-occurrence matrix and its statistical features as a new approach for face Recognition," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 19, no. 1, 2011, doi: 10.3906/elk-0906-27.

[28] Haralick, R.M and K. Shanmugam, "Texture features for image classification," *EEETrans. Syst. Man Cybern. SMC*, vol. 3, no. 6, pp. 610–621, 1973.

[29] M. Huang, H. Shu, Y. Ma, and Q. Gong, "Content-based image retrieval technology using multi-feature fusion," *Opt. - Int. J. Light Electron Opt.*, vol. 126, no. 19, pp. 2144–2148, 2015, doi: 10.1016/j.ijleo.2015.05.095.

[30] O. I. Abiodun *et al.*, "Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition," *IEEE Access*, vol. 7, pp. 158820–158846, 2019, doi: 10.1109/ACCESS.2019.2945545.

[31] V. Ojha, A. Abraham, and V. Snášel, "Metaheuristic design of feedforward neural networks: a review of two decades of research," *Eng Appl Artif Intell*, vol. 60, pp. 97–116, 2017.

[32] A. A. Heidari, H. Faris, I. Aljarah, and S. Mirjalili, "An efficient hybrid multilayer perceptron neural network with grasshopper optimization," *Soft Comput.*, vol. 23, no. 17, pp. 7941–7958, 2019, doi: 10.1007/s00500-018-3424-2.