ICIC Express Letters
Part B: Applications
Volume **14**, Number **5**, May **2023**

ICIC International ©2023 ISSN 2185-2766
pp. 457–465

IMPROVING THE EFFICIENCY OF IMAGE RETRIEVAL BASED ON A MODEL COMBINING GRAPH AND SOM

Nguyen Thi Uyen Nhi[1], Nguyen Minh Hai[2,3,*] and Van The Thanh[4]

[1]Faculty of Statistics and Informatics
University of Economics
The University of Danang
No. 71, Ngu Hanh Son St., Bac My An Ward, Ngu Hanh Son Dist., Da Nang City 556920, Vietnam
nhintu@due.edu.vn

[2]Graduate University of Science and Technology
Vietnam Academy of Science and Technology, Vietnam
No. 18, Hoang Quoc Viet St., Nghia Do Ward, Cau Giay Dist., Ha Noi 122000, Vietnam
minhhaidhsp@gmail.com

[3]Faculty of Physics
[4]Faculty of Information Technology
Ho Chi Minh City University of Education
No. 280, An Duong Vuong St., Ward 4, Dist. 5, Ho Chi Minh City 749000, Vietnam
thanhvt@hcmue.edu.vn
[*]Corresponding author: hainm@hcmue.edu.vn

Received November 2022; accepted January 2023

Abstract. *Image retrieval is a research issue that has been of interest in recent times. In our previous work, we constructed a neighbor graph (called Graph-CTree) for storing and retrieving large image data. In this paper, we propose methods to improve the efficiency of image retrieval: creating a SgC-Tree model from a combination of the neighbor graph (Graph-CTree) and self-organizing map (SOM). In this paper, SOM is assembled from clusters of Graph-CTree graphs, called grSOM, with input weight vectors taken during training C-Tree. So, the weights do not have to be tweaked too much during training, so the training time of grSOM is faster. grSOM is more flexible and allows scaling after training. Content-based image retrieval system has been built on SgC-Tree, called CBIR-SgC, and experimented on COREL and ImageCLEF datasets to evaluate the effectiveness and correctness of the proposal.*
Keywords: Content-based image retrieval, SOM, grSOM, SgC-Tree, CBIR-SgC

1. **Introduction.** Digital image data plays an important role in many modern systems serving the needs of use, exploitation, and discovery. The influence of the image database depends on the content and extent of information sharing for each image dataset. Therefore, the techniques of content-based retrieval are essential for various image retrieval applications.

Content-based image retrieval (CBIR) focuses on the study of techniques to extract low-level features and the structures to store these data [1,2]. CBIR extracted low-level visual features [2-4] to identify objects that are focused on in the image. There were many techniques of image retrieval that have been widely applied in many different digital systems, such as techniques to build a model for extraction low-level features [1,2], techniques for classification data [8,9], and content-based image retrieval for hierarchical databases [1,11]. In 2018, Cevikalp et al. [10] proposed a method for large-scale image retrieval by using binary hierarchical trees and transductive support vector machines (TSVM). However, this structure was only used to classify images but has not yet created a storage

DOI: 10.24507/icicelb.14.05.457

structure for the images. Jiu and Sahbi [11] used deep multi-layer networks based on nonlinear activation functions for image annotation. The support vector machines (SVM) technique is applied to layering images at the output layer to extract a semantic level according to visual information for similar pocket-based images from Bag-of-Words. In this method, a deep multi-layer network is fixed the number of layers, so the image classification is limited. Kanwal et al. [12] proposed to use the combined feature from the 34-layer ResNet architecture together with the PCA-reduced (principal component analysis) selective feature extraction technique. The convolutional neural network (CNN) is then used in conjunction with the BoW visual bag of words to index, rank, and retrieve the classified images. This method gives high accuracy in image search (0.89). However, the computational complexity of this method is high, and the query time is long. Dhingra and Bansal [1] used a texture feature extraction method that combines gray level co-occurrence matrix (GLCM), discrete wavelet transforms (DWT), Gabor transforms, and local binary patterns (LBP). Then a cascade forward back propagation neural network (CFBPNN) is used as a classifier after the feature extraction step to increase the average accuracy of the system. Although this method only extracts texture features, with the use of neural networks for image classification, the accuracy is high; however, the average image search time is higher.

From the shortages of the above studies, it can be seen that the data organization methods will affect the efficiency of the image retrieval system. Therefore, building a good storage structure will improve the accuracy of image retrieval. This is the impetus for our research. In our previous papers, a neighbor graph named Graph-CTree [5] has been built to improve the efficiency of the content-based image retrieval with C-Tree [6]. From the most relevant leaf node searched on the C-Tree to the input query image, a neighbor graph of clusters is generated from the nodes associated with this leaf node, improving search efficiency. Our proposed methods previously improved the precision of image retrieval. However, the cluster selection criterion of Graph-CTree is based on distance measure, so it can lead to errors if the amount of data is large. The SOM (self-organizing maps) [7] network overcomes these cluster graph problems because of the winning cluster selection criteria. Our proposals in this paper include: 1) Building a grSOM is assembled from clusters on the Graph-CTree neighbor graph and a set of weight vectors trained from the C-Tree; 2) Building a content-based image retrieval CBIR-SgC is a combination of Graph-CTree structure and grSOM.

The remainder of the paper includes the following. Section 2 describes improvements in a combination model SgC-Tree. Section 3 presents a CBIR model based on SgC-Tree. Section 4 builds an application of the CBIR-SgC system based on the proposed model. The experiment was conducted on COREL, ImageCLEF datasets to compare with other methods, thereby proving the effectiveness of the proposed system. Section 5 presents conclusions of the paper.

2. A Combination Model SgC-Tree.

2.1. A general description of C-Tree and Graph-CTree.

In our own previous studies, a balanced clustering tree (C-Tree) [6] is created based on the K-means method for the feature vectors of image datasets. However, when splitting a node, the C-Tree can have similar elements but still split into two separate nodes. In the worst case, these elements are in two different branches. Therefore, similar image retrieval will not find branched similar elements. To overcome the disadvantages of C-Tree, a neighbor graph (Graph-CTree) [5] was built from leaf nodes on the C-Tree. The Graph-CTree has solved most of C-Tree's problems, improving image retrieval performance. However, the criterion for selecting the cluster of the graph is by similar measure, which may lead to errors if the tree performs splitting many nodes and a large number of layers. So, it is necessary to add the criterion that selects the winning leaf node according to the weights of the elements that appear the most at each leaf node. Therefore, the paper proposes SOM to find the winning cluster according to the representative classification.

2.2. **SgC-Tree structure.** SgC-Tree structure is a combination of C-Tree, Graph-CTree neighbor cluster graph, and self-organizing maps (SOM) network. The SOM training process is the weight training process [7]. Adjusting the weights will make the SOM achieve the best clustering requirement. However, the weight adjustment process takes a lot of time with large input images and random initialization, so each training can generate completely different maps. At the same time, the SOM is static after training, so when adding new data the map will misclassify the input data, so the SOM must be trained from the beginning.

Therefore, a self-assembled network from the graph is proposed in the paper to overcome the disadvantages from SOM. This SOM network is assembled from clusters of Graph-CTree graphs, called grSOM, with input weight vectors taken during training C-Tree with the following advantages.

- A set of input weight vectors of the grSOM network is trained on a C-Tree. Initially, grSOM has a stable weight vector with high accuracy, the weights do not have to be tweaked too much during training, so the training time of grSOM is faster than that of the traditional SOM.
- The grSOM network is more flexible and allows scaling after training. The grSOM network is assembled each leaf node cluster of the Graph-CTree graph, so if a new leaf node is generated, it will be trained on the tree with its own weight without having to train from the beginning of the whole network.

Figure 1 shows the structure of SgC-Tree: C-Tree – Graph-CTree – grSOM. Figure 1 shows that a self-organizing map named grSOM has clusters assembled from the neighbor clustering graph Graph-CTree, with the set of input weight vector W obtained from the building C-Tree. The structure of the grSOM inherits the structure of the SOM and the Graph-CTree. The SOM is a transmission neural map that uses unsupervised learning algorithms known as competitive learning and "self-organization". It sorts output for a geometrical or spatial representation of the original data. The training process of the grSOM is based on competitive learning to select the winning cluster according to the representative classification.

The grSOM is an SOM whose inputs are the feature vectors of the image $f = (f_1, f_2, \dots, f_m)$, in which, each vector f_i has n dimensions $f_i = (v_1, v_2, \dots, v_n)$, $f_i \in \{0, 1\}$ and the output layer consists of components assembled from the clusters of leaf nodes in the Graph-CTree. The input and output stages are fully connected by weight vectors $W_i = (w_1, w_2, \dots, w_n)$, $w_i \in \{0, 1\}$. Thus, the weight vectors p set is $W = (W_1, W_2, \dots, W_p)$.

Figure 2 shows an example of assembling leaf nodes from Graph-CTree to grSOM. The training process on the grSOM includes the following steps:

- Step 1. To assemble leaf nodes from Graph-CTree to grSOM;
- Step 2. To initialize the initial weight from the weight gained during training on the C-Tree;
- Step 3. To select randomly a vector f_i as the model for training;
- Step 4. To find the winning clusters based on the Sigmoid function;
- Step 5. To refine the weights based on the generalized reduced gradient (GRG) method;
- Step 6. To take the sample for the next training and to repeat Step 3 until the algorithm is optimized.

3. **Model of Content-Based Image Retrieval CBIR-SgC.** The content-based image retrieval CBIR-SgC system is a combination of Graph-CTree and SOM. Figure 3 shows a

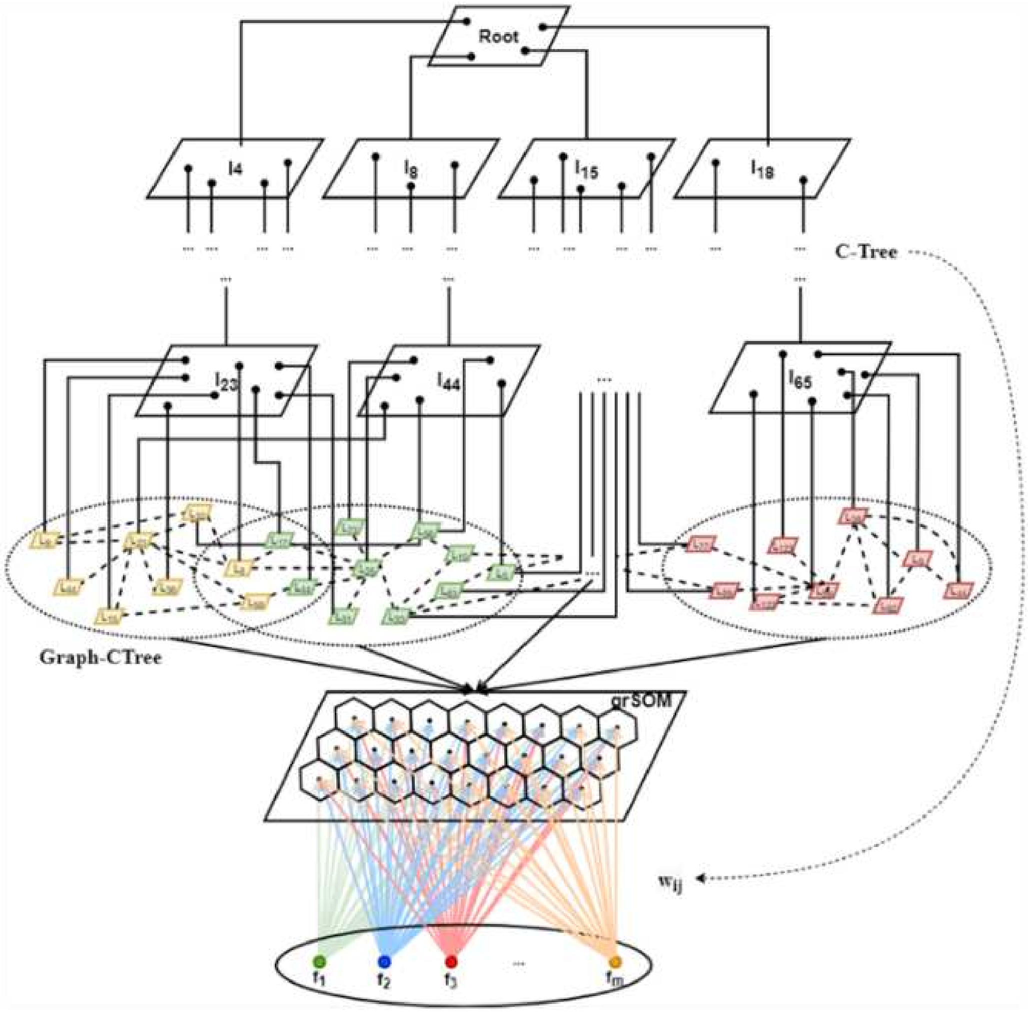


FIGURE 1. Structure of SgC-Tree

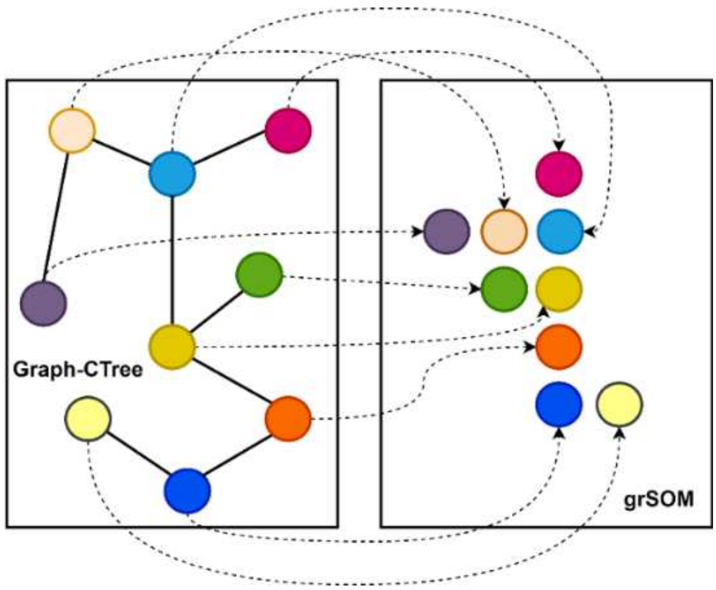


FIGURE 2. Example of assembly of a leaf node from Graph-CTree to grSOM

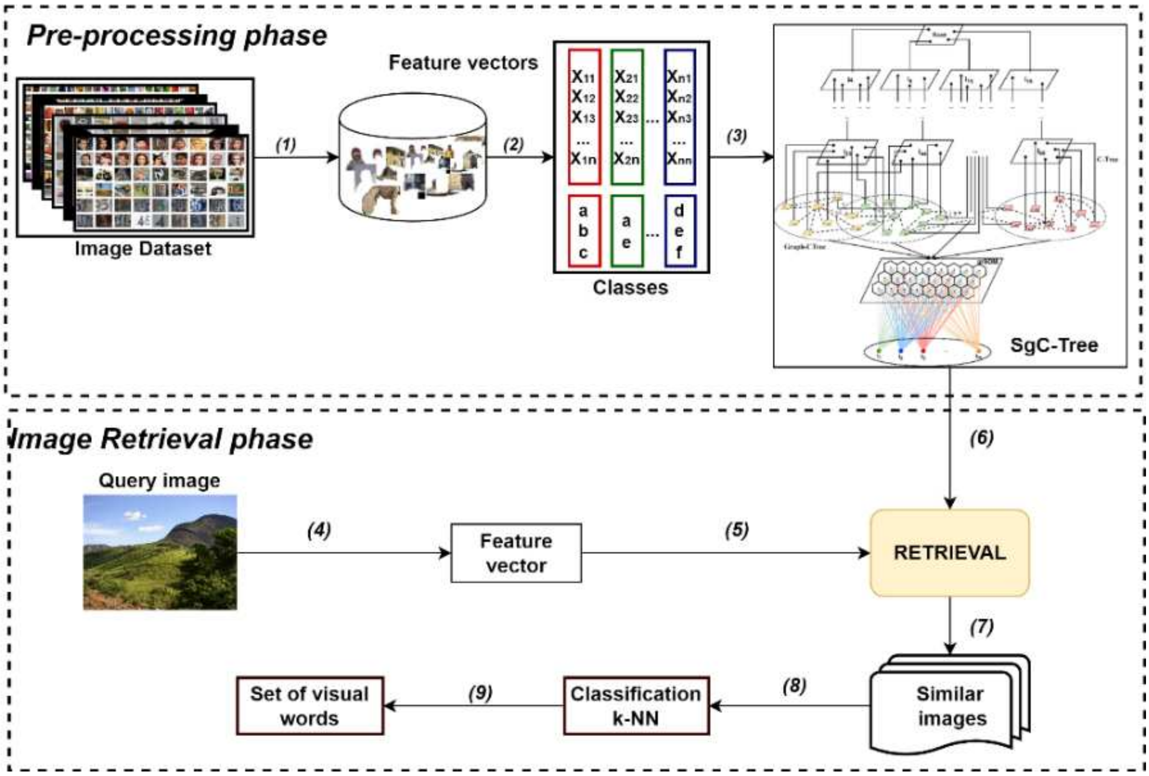


FIGURE 3. Model of content-based image retrieval based on SgC-Tree

model of a content retrieval system based on SgC-Tree, consisting of two specific phases as follows:

(1) Preprocessing phase:

Step 1. From the image dataset (1), low-level feature vectors f are extracted (2);

Step 2. Set of trained weight vectors serves as initial weights for the SOM map assembled from the graph (3);

(2) Query phase:

Step 1. For an input query image, the low-level feature vector is extracted (4);

Step 2. From the feature vector (5), query on SgC-Tree (6): To search on C-Tree to find the most suitable leaf node, then to search the neighbor clusters of that leaf node on the neighbor graph. At the same time, to search on grSOM for the winning cluster, and to take neighbors for the winning cluster. To assign similar image sets to find the best one (7);

Step 3. The k-NN algorithm (8) is performed on the set of similar images to find the set of visual words (9).

This querying procedure shows that the final set of similar images found is the result of the intersection of the query on C-Tree, Graph-CTree, and grSOM. Therefore, the image query efficiency is the best of the proposed models. The result of this querying is a set of similar images and a set of visual words.

4. Building an Application of the Content-Based Image Retrieval.

4.1. **Experimental environment.** CBIR-SgC system is built to retrieve images with SgC-Tree. CBIR-SgC is implemented based on dotNET Framework 4.8 platform and C # programming language. The graphs are built on Mathlab 2015. The configuration of the computer in the experiment: Intel (R) CoreTM i7-8750H, 2.70GHz CPU, 8GB of RAM, and Windows 10 Professional operating system. The datasets used in the experiment are image datasets such as COREL (1,000 images) and ImageCLEF (20,000 images).

4.2. **Experimental application.** For an input image, a feature vector is extracted and retrieved on SgC-Tree to find a set of similar images. Then, k-NN algorithm is performed on the set of similar images to find the set of visual words. Figure 4 shows an example of an image retrieval process based on SgC-Tree.

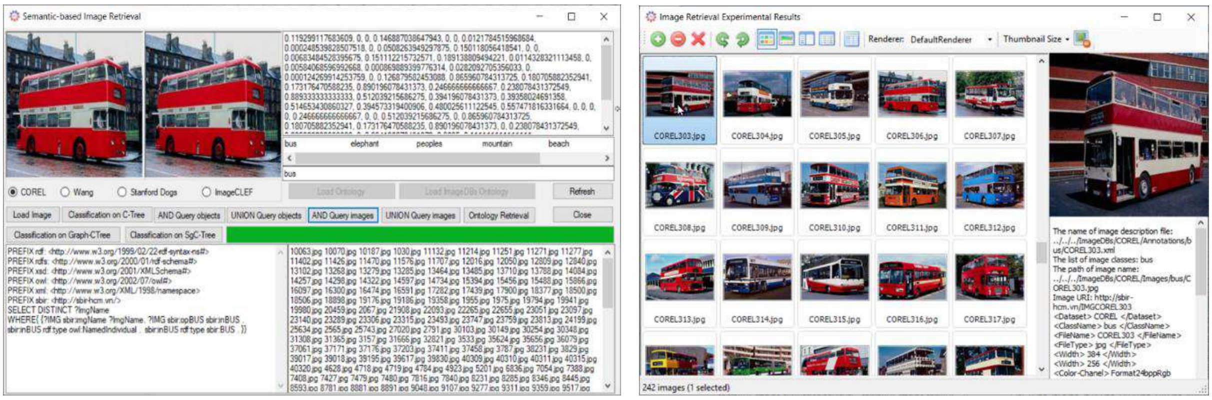


FIGURE 4. A result of the content-based image retrieval of SgC-Tree (CBIR-SgC)

4.3. **Experimental evaluation.** To evaluate the effectiveness of image retrieval, this paper uses the following factors for evaluation like precision, recall and F-measure, query time (milliseconds). The evaluations are implemented on COREL, ImageCLEF datasets in Table 1 and Table 2.

TABLE 1. Performance of retrieval systems on the COREL dataset

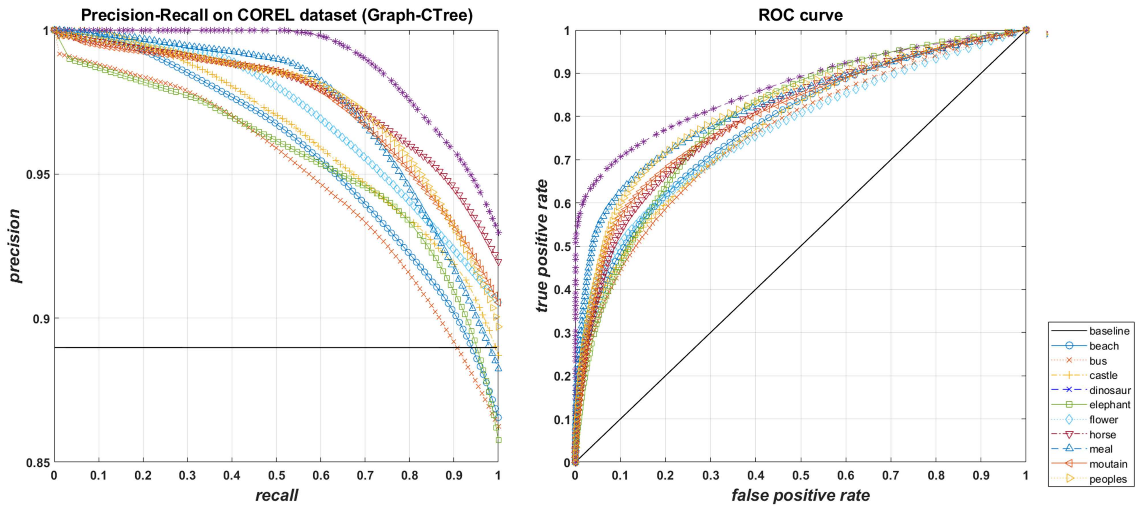
Performance	**Precision**	**Recall**	**F-measure**	**Query time (ms)**
Graph-CTree	0.888473	0.884555	0.886346	72.65352
SgC-Tree	**0.913212**	**0.923649**	**0.9183137**	**86.1635**

TABLE 2. Performance of retrieval systems on the ImageCLEF dataset

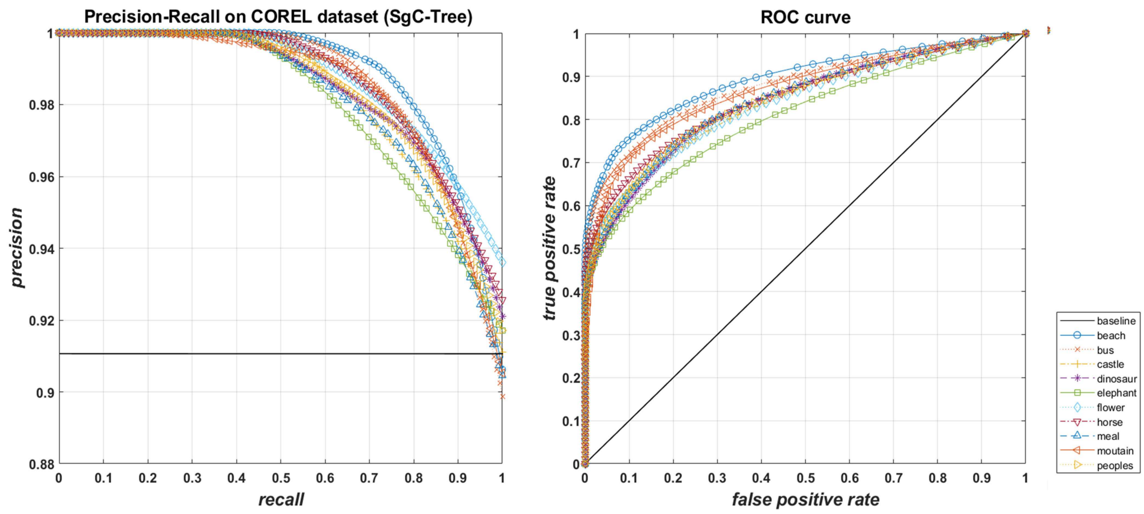
Performance	**Precision**	**Recall**	**F-measure**	**Query time (ms)**
Graph-CTree	0.839814365	0.78073583	0.806435717	239.9458
SgC-Tree	**0.874402**	**0.864789**	**0.869484**	**242.1663**

Performance comparison of image retrieval system on SgC-Tree (CBIR-SgC) shows that the precision is better than that of image retrieval system on Graph-CTree. In addition, Precision-Recall and ROC curve graphs [2] were performed to evaluate the accuracy of image retrieval system based on SgC-Tree compared with Graph-CTree (Figure 5 and Figure 6). The area under the AUC curve (area under the curve) and the spatial limitation of the ROC is a measure of the precision of the query. The larger AUC is, the higher the precision is. Besides, in the ROC curve graph, a baseline diagonal divides the space of the ROC into two parts. The points above the diagonal represent the correct classification results. The points below the diagonal are the results of false classification. Points further from the baseline will give better classification results than points located near the baseline.

The performance of the image retrieval system in the paper is compared with modern methods from other studies on the same dataset. Table 3 and Table 4 show the results of comparing the mean average precision (MAP) of the image retrieval methods on the COREL, ImageCLEF datasets, respectively.



(a)

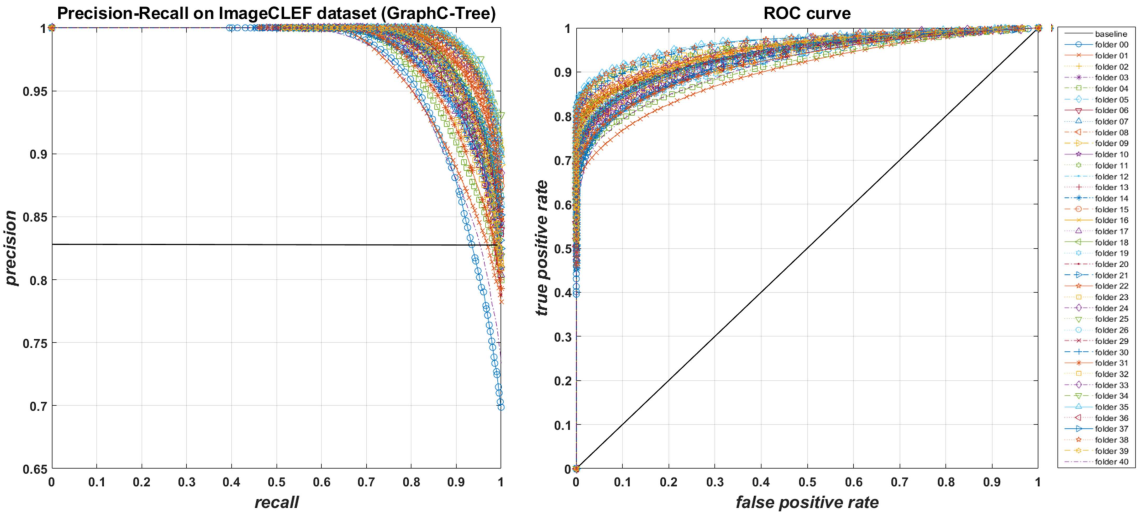


(b)

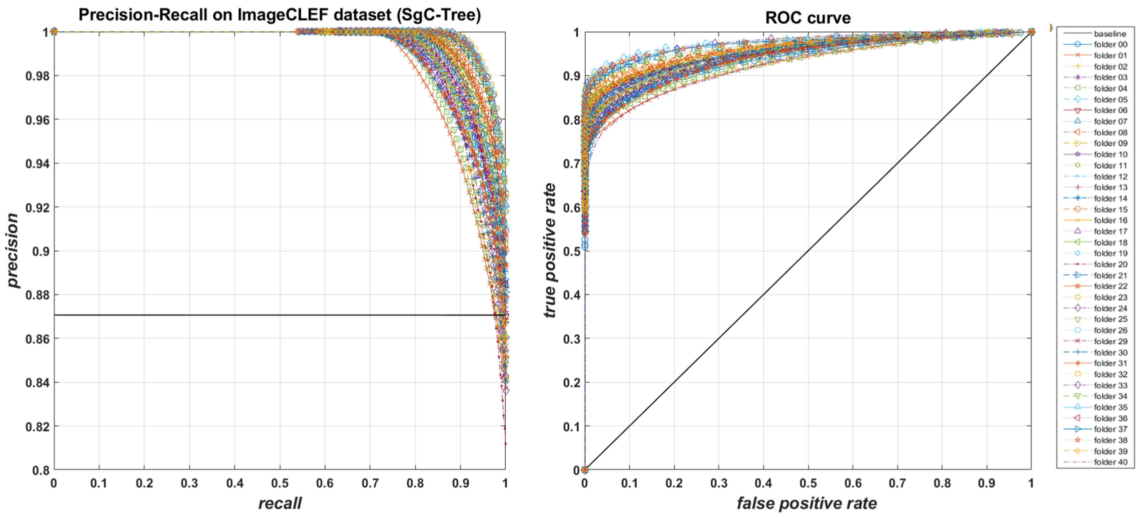
FIGURE 5. Performance on Graph-CTree (a), and SgC-Tree (b) of COREL dataset

The data in Table 3 and Table 4 show that the proposed method has higher precision compared with other image retrieval methods such as using deep learning techniques [1,12,13], and hash functions [5], on the same dataset. The proposed method has better accuracy than other methods because this result is combined to retrieve images on three different structures (C-Tree, Graph-CTree, SgC-Tree), in which these structures overcome each other's weaknesses. This shows that our proposed method is effective in solving the content-based image retrieval problem and semantic analysis for single-object images (COREL), and multi-object images (ImageCLEF). So, the methods proposed in this paper improve the precision of content-based image retrieval on Graph-CTree. These suggestions are correct and effective.

5. **Conclusions.** This paper has proposed methods to improve a neighbor graph Graph-CTree that was built in our previous own study. First, a model combining grSOM and Graph-CTree, called SgC-Tree, is created to improve the efficiency of image retrieval. The SgC-Tree model adds the winning leaf node selection criterion, making the clustering method better, and the precision of image retrieval higher. The experiment is performed on image datasets, such as COREL (1,000 images), ImageCLEF (20,000 images). CBIR-SgC system has superior precision compared to our previous proposals. Experimental



(a)



(b)

FIGURE 6. Performance on Graph-CTree (a), and SgC-Tree (b) of ImageCLEF dataset

TABLE 3. Comparison MAP of methods on COREL dataset

Methods	MAP
Multi-feature with neural network [3]	0.7941
Fusion feature ResNet-34 + PCA + CNN [12]	0.89
Multi-feature and SVM [2]	0.7657
Texture features + CFBPNN [1]	0.82
SgC-Tree	**0.91321**

TABLE 4. Comparison MAP of methods on ImageCLEF dataset

Method	MAP
Fusion hashing network + binary code matrix + CNN [13]	0.8038
Hybrid deep learning architecture [14]	0.797
Consistency Preserving Adversarial Hashing [15]	0.8324
SgC-Tree	**0.8744**

performance is compared with other methods on the same image dataset to evaluate the proposed model, method, and algorithm. The comparison results show that the query system CBIR-SgC is more accurate than other studies on the same image dataset. This shows that our proposals in this paper are effective and correct. In future studies, we will apply deep learning techniques and combine them with content image retrieval methods to improve image retrieval efficiency.

Acknowledgment. The authors would like to thank The University of Danang, University of Economics, and Ho Chi Minh City University of Education, which are sponsors of this research. This research is part of University-level research project granted by The University of Danang, University of Economics with grant number of T2022-04-21. This research is funded by Ho Chi Minh City University of Education Foundation for Science and Technology under grant number CS.2022.19.55. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] S. Dhingra and P. Bansal, Relative examination of texture feature extraction techniques in image retrieval systems by employing neural network: An experimental review, *Proceedings of International Conference on Artificial Intelligence and Applications*, Springer, Singapore, pp.337-349, 2021.

[2] M. Garg and G. Dhiman, A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants, *Neural Computing and Applications*, vol.33, pp.1311-1328, 2021.

[3] R. Raja, S. Kumar and M. R. Mahmood, Color object detection-based image retrieval using ROI segmentation with multi-feature method, *Wireless Personal Communications*, vol.112, no.1, pp.169-192, 2020.

[4] L. Xie, J. Wang, B. Zhang and Q. Tian, Fine-grained image search, *IEEE Transactions on Multimedia*, vol.17, no.5, pp.636-647, 2015.

[5] N. T. U. Nhi and T. M. Le, A model of semantic-based image retrieval using C-Tree and neighbor graph, *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol.18, no.1, pp.1-23, 2022.

[6] N. T. U. Nhi and T. M. Le, Semantic-based image retrieval using balanced clustering tree, *World Conference on Information Systems and Technologies*, Springer, Cham, pp.416-427, 2021.

[7] C. S. Wickramasinghe, K. Amarasinghe and M. Manic, Parallalizable deep self-organizing maps for image classification, *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp.1-7, 2017.

[8] M. Garg, H. Singh and M. Malhotra, Fuzzy-NN approach with statistical features for description and classification of efficient image retrieval, *Modern Physics Letters A*, vol.34, no.3, 1950022, 2019.

[9] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie and M. Li, Bag of tricks for image classification with convolutional neural networks, *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.558-567, 2019.

[10] H. Cevikalp, M. Elmas and S. Ozkan, Large-scale image retrieval using transductive support vector machines, *Computer Vision and Image Understanding*, vol.173, pp.2-12, 2018.

[11] M. Jiu and H. Sahbi, Nonlinear deep kernel learning for image annotation, *IEEE Transactions on Image Processing*, vol.26, no.4, pp.1820-1832, 2017.

[12] K. Kanwal, K. T. Ahmad, R. Khan, A. T. Abbasi and J. Li, Deep learning using symmetry, fast scores, shape-based filtering and spatial mapping integrated with CNN for large scale image retrieval, *Symmetry*, vol.12, no.4, 612, 2020.

[13] V. E. Liong, J. Lu, Y. P. Tan and J. Zhou, Cross-modal deep variational hashing, *Proc. of the IEEE International Conference on Computer Vision*, pp.4077-4085, 2017.

[14] K. S. Arun and V. K. Govindan, A hybrid deep learning architecture for latent topic-based image retrieval, *Data Science and Engineering*, vol.3, no.2, pp.166-195, 2018.

[15] A. Yang, X. Yang, W. Wu, H. Liu and Y. Zhuansun, Research on feature extraction of tumor image based on convolutional neural network, *IEEE Access*, vol.7, pp.24204-24213, 2019.