MINISTRY OF EDUCATION AND TRAINING

VIETNAM ACADEMY OF SCIENCE AND TECHNOLOGY

GRADUATE UNIVERSITY OF SCIENCE AND TECHNOLOGY

NGUYEN MINH HAI

DEVELOPING A SEMANTIC-BASED IMAGE RETRIEVAL ALGORITHM

SUMMARY OF DISSERTATION ON COMPUTER

Major: Computer Science

Code: 948 01 01

Ho Chi Minh City - 2024

The dissertation is completed at: Graduate University of Science and Technology, Vietnam Academy Science and Technology

Supervisors:

1. Supervisor 1: Assoc. Prof. Dr. Tran Van Lang, Institute of Applied Mechanics and Informatics, Vietnam Academy of Science and Technology
2. Supervisor 2: Dr. Van The Thanh, Ho Chi Minh City University of Education

Referee 1: Assoc. Prof. Dr. Tran Cong Hung

Referee 2: Assoc. Prof. Dr. Huynh Tuong Nguyen

Referee 3: Dr. Le Quang Minh

The dissertation is examined by Examination Board of Institute of Applied Mechanics and Informatics, Vietnam Academy of Science and Technology at 09:00, December 6, 2024.

The dissertation can be found at:

1. Graduate University of Science and Technology Library
2. National Library of Vietnam

INTRODUCTION

1. Overview

Machine vision focuses mainly on the application of electronic and computer-aided imaging, especially in the analysis of images collected from systems such as pedestrian tracking, traffic analysis, security surveillance, and deep space exploration. However, to achieve better results in machine vision applications, people are increasingly promoting the combination of machine learning with image processing methods in the analysis process. This combination is considered a way to create more intelligent systems and has become popular in image retrieval applications.

2. Urgency of the thesis

The problem of semantic image retrieval is a problem that is of great concern to researchers around the world, and has good performance when solved. With the desire to contribute an effective semantic image retrieval method, the thesis carried out the topic: "Developing a semantic-based image retrieval algorithm".

3. Research Objectives of the Thesis

The research objectives of the thesis focus on the following issues:

(1) Analyze and extract semantic features from images.

(2) Understand and process image components to create a semantic image.

(3) Develop an image search and classification system that can understand and respond to user requests effectively.

4. Research Methodology

Theoretical method: Synthesize works related to semantic querying using machine learning methods and tree-like storage structures.

Experimental method: Based on the methods and models proposed in the thesis, implement experimental program installation on popular data sets.

5. Research Subjects and Scope

Research objects: Search algorithms; semantics in images; image data.

Scope of study: Identify certain semantic types such as objects, actions or contexts.

6. Thesis Contributions

Developing a semantic image search algorithm based on visual features of images based on the GP-Tree data structure to improve image search accuracy, including:

- Building a GP-Tree hierarchical clustering data structure to organize the storage of image feature vectors
- Developing a GP-Tree structure based on supervised and semi-supervised learning algorithms to increase image search efficiency
- Building a semantic image search system based on the GP-Tree structure and ontology to demonstrate the image search efficiency of the proposed methods

7. Thesis Content and Structure

The structure of the thesis includes:

Chapter 1: Overview of image search

This chapter presents an overview of the image search problem, focusing on two main directions: content-based image search (CBIR) and semantic image search (SBIR). Related studies are surveyed and analyzed to identify challenges and limitations in existing methods. From there, the thesis proposes research directions to overcome these limitations. This section also describes in detail how to organize experiments, including environment setup, data set selection and use, and criteria for evaluating search performance.

Chapter 2: GP-Tree structure for image storage and retrieval

This chapter focuses on using tree structure to store and index image data sets. The GP-Tree structure is described in detail, including operations for adding, editing, splitting, and deleting elements. A semantic image search model based on GP-Tree is proposed to improve search performance. Experiments on popular data sets such as Wang, MS-COCO, and ImageCLEF demonstrate the effectiveness of the model.

Chapter 3: GP-Tree Improvement with Graph-GPTree for Image Search

This chapter presents a method to improve GP-Tree by combining it with the neighboring cluster graph, forming a Graph-GPTree structure. This improvement aims to improve the efficiency of storing and searching similar elements. An image search model based on Graph-GPTree is proposed, with the goal of improving the accuracy and speed of searching. Experiments are conducted on the Wang, MS-COCO, and ImageCLEF datasets to evaluate the effectiveness of the method.

Chapter 4: SgGP-Tree Network Structure for Semantic-Based Image Search

This chapter proposes a new improvement, combining Graph-GPTree and self-organizing network (SOM), called SgGP-Tree. This structure is designed to improve the efficiency of storing and searching images. At the same time, a semantic image search model based on ontology is proposed, called SBIR-GP. Experiments on standard datasets such as Wang, MS-COCO, and ImageCLEF have confirmed the superior performance of the method.

Conclusion and future development direction: Present the results achieved and the direction for further development of the thesis.

List of works of the author: List the works that the author has published during the thesis.

References: List the documents that the thesis has referred to.

CHAPTER 1. OVERVIEW OF IMAGE RETRIEVAL

1.1. Overview of image retrieval

The image retrieval problem presented in this thesis is defined as finding the set of images with the closest similarity to the input image based on the similarity measure between the images. Figure 1.1 describes the methods in image retrieval

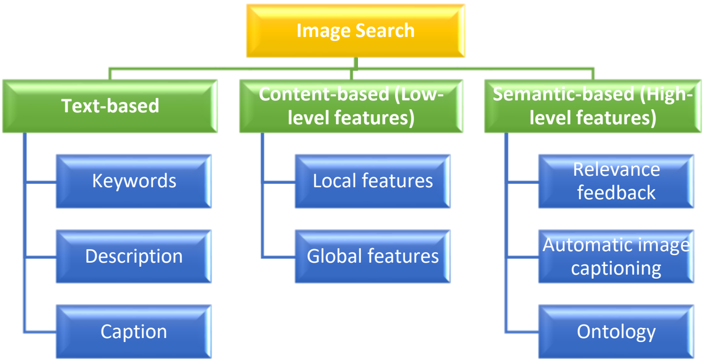

Figure 1.1. Types of image retrieval.

1.2. Common features in image retrieval

A feature is defined as the identification of a visual attribute of an image. In general, image features can be global or local.

1.3. Related research works on image retrieval

1.3.1. Text-based image retrieval

The text-based method is a simple traditional keyword-based search method (**Figure 1.2**). Images are indexed by content, such as image caption; file name, web page title, and alt tags.... and stored in a database.

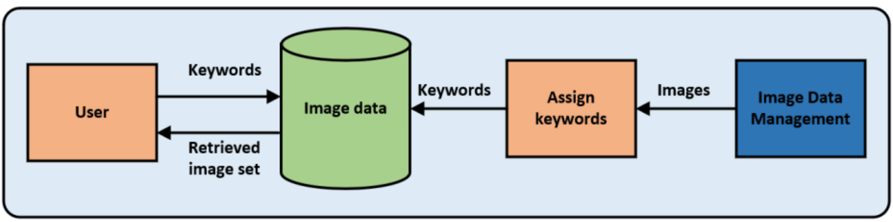

Figure 1.2. Text-based Image Retrieval System

1.3.2. Content-Based Image Retrieval

In a content-based image retrieval system, the efficiency of the computation depends on the ability to extract low-level features and measure similarity. In the content-based image retrieval method (**Figure 1.3**), low-level visual features such as color, shape, texture, and spatial layout are used

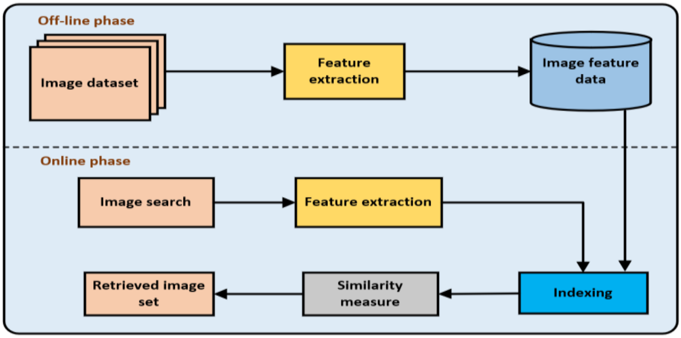

Figure 1.3. Content-based Image Retrieval System

1.3.3. Semantic-Based Image Retrieval

In the field of image retrieval, the main challenge is to convert images into low-level features that can be computed by computers and associate them with high-level concepts to reduce the semantic gap. Many semantic image retrieval methods have been proposed to reduce the semantic gap such as: machine learning-based semantic image retrieval, ontology-based semantic image retrieval. The semantic image retrieval system is depicted in **Fig 1.4**

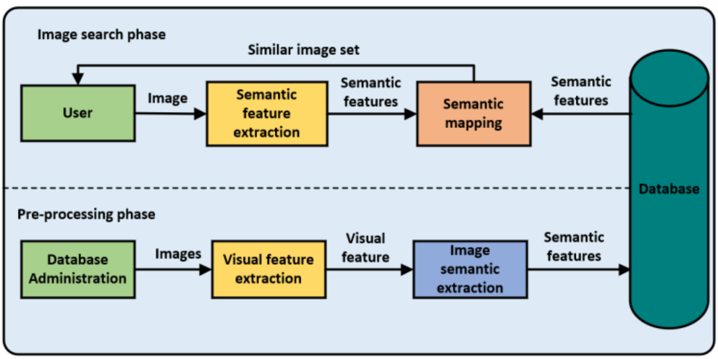

Figure 1.4. Semantic-based image retrieval system

1.4. Experimental setup and evaluation methods

The proposed models are built and tested based on the dotNet Framework 4.5 and Python 3.7 platforms

The datasets used in the thesis experiments are popular datasets and are widely used in research works on image retrieval, including: Wang, ImageCLEF, MS-COCO.

The results that need to be evaluated for performance include: Image classification results and image search accuracy. The performance evaluation values of image classification and search used in the thesis include P (precision), coverage R (recall) and compromise F_m (F-measure).

1.5. Chapter conclusion

This chapter provides a detailed overview of modern image retrieval methods, including two main approaches: content-based image retrieval and semantic image retrieval.

This chapter also presents specific experimental organization methods, including the setup of the experimental environment, the selection and preparation of experimental data sets, as well as the performance evaluation metrics of the search methods. The next chapter will present a new data structure, the hierarchical clustering tree, which is proposed to optimize the storage and indexing of image data sets. This structure not only supports the fast retrieval process but also enhances the efficiency of semantic image search, allowing the system to process large image data sets quickly and efficiently.

CHAPTER 2. GP-TREE STRUCTURE FOR IMAGE STORAGE AND RETRIEVE

This chapter focuses on using tree structure to store and index image datasets. The GP-Tree structure is described in detail, including operations of adding, editing, splitting, and deleting elements. A semantic image search model based on GP-Tree is proposed to improve the search performance. Experiments on popular datasets such as Wang, MS-COCO, and ImageCLEF demonstrate the effectiveness of the model. The content of this chapter is directly related to two published works [CT4] and [CT5]; and also indirectly related to works [CT1], [CT2], [CT3].

2.1. GP-Tree Data Structure

Based on the multi-branch tree structure and the K-Means clustering method, the GP-Tree structure is constructed by splitting the leaf node into two nodes if the number of elements at that leaf node exceeds a given value M, and at the same time, each node in the child nodes can be formed based on a threshold θ, which is used to evaluate the similarity of the data. Therefore, the GP-Tree grows in the leaf direction and develops into a multi-branch tree in which each leaf node is a data cluster consisting of similar elements. The data elements are feature vectors for each image and are stored in turn on the GP-Tree, from which operations on the tree are performed. The structure of the GP-Tree hierarchical clustering tree is illustrated in **Figure 2.1**

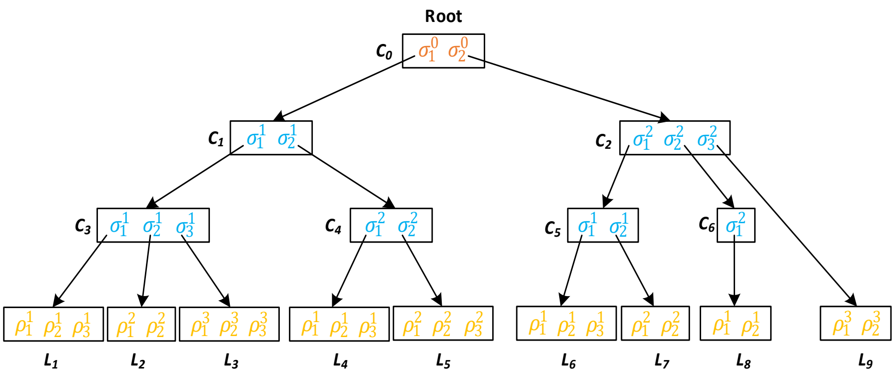

Figure 2.1. Hierarchical clustering GP-Tree with 3 levels

2.2. Structure image retrieval based on GP-Tree

2.2.1. GP-Tree based image retrieval system

The image retrieval system using GP-Tree consists of two phases: (1) The preprocessing phase performs segmentation, feature extraction from the image dataset, extracts concept classes, and organizes storage on the GP-Tree; (2) The query phase searches for similar images by content on the GP-Tree and extracts the image semantics. **Figure 2.2** presents the image search system model based on GP-Tree, with two specific phases as follows:

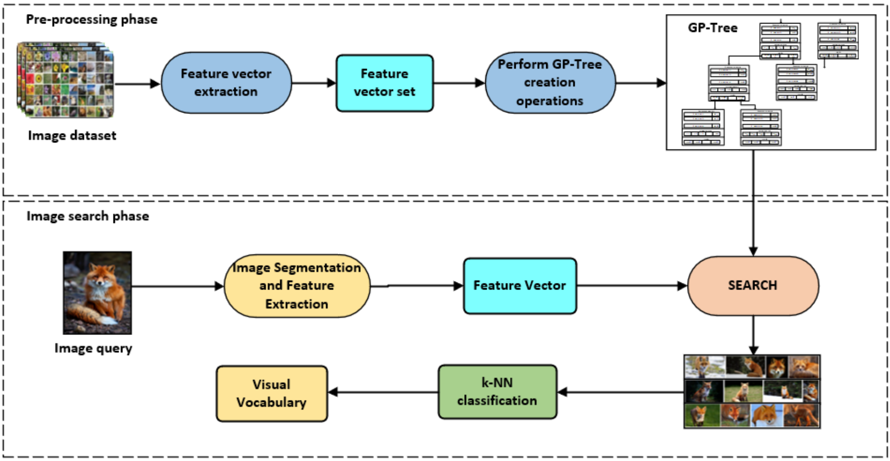

Figure 2.2. GP-Tree-based image retrieval system model

2.2.2. GP-Tree creation algorithms

The algorithms used in constructing the GP-Tree structure include:

Algorithm 2.1: Add data elements **Input:** ρ,η,θ,"GP-Tree" **Output:** GP-Tree after adding leaf node elements **Function:** insertED(ρ,η,θ,"GP-Tree")	**Algorithm 2.2:** Split a node on the GP-Tree **Input:** L_s,"GP-Tree" , M **Output:** GP-Tree after splitting **Function:** splitLeafNode(GP-Tree, L_s, M)

Algorithm 2.3: Delete an element of a leaf node on the tree **Input:** GP-Tree, L_s, ρ **Output:** GP-Tree after deleting leaf node elements. **Function:** deleteLeafElement(GP-Tree, L_s, ρ)	**Algorithm 2.4:** Delete an element of an internal node on the tree **Input:** GP-Tree, C_s, σ **Output:** GP-Tree after deleting the element of the internal node. **Function:** deleteInternalNodeElement(GP-Tree, C_s, σ)
Algorithm 2.5: Create GP-Tree **Input:** Image dataset Γ, threshold θ **Output:** GP-Tree **Function:** createGPT(Γ,θ)	**Algorithm 2.6:** Create visual vocabulary **Input:** Similar image set Ω, threshold γ **Output:** Visual vocabulary set W **Function:** CreateVW(Ω,γ)

2.2.3. Experiment and evaluation of results

Experiments were performed on WANG, ImageCLEF and MS-COCO image sets, each with different characteristics and quantities. The experimental results of GP-Tree are shown in **Table 2.1**, with the parameters M (maximum number of elements of leaf nodes) and N (maximum number of elements of internal nodes) adjusted depending on each image set. **Table 2.1** also provides the number of image samples with the best precision (top precision – P@). The image retrieval system is based on the GP-Tree structure and the search results are depicted in **Figure 2.3**.

Table 2.1: GP-Tree experiment results

Image Dataset	Number of Images	Experimental Parameters		Experimental Time (seconds)	Number of Leaf Clusters	Number of Internal Clusters	Number of P@ Samples	P@ Rate
		M	N					
WANG	10,800	100	40	17,839.47	218	32	2,240	20%
ImageCLEF	20,000	150	50	32,173.95	432	67	6,000	20%
MS-COCO	163,957	300	70	158,765.84	782	314	44,188	20%

To evaluate the image retrieval performance, factors such as Precision, Coverage, Relevance and Search Time (milliseconds) are used, with the indices and average search times of image directories on the WANG, ImageCLEF and MS-COCO datasets summarized in **Table 2.2.**

Table 2.2: Image retrieval performance of the GP-Tree based image retrieval system above.

Image Dataset	Evaluation Metrics			
	Accuracy	**Recall**	**F-measure**	**Average Search Time (ms)**
WANG	0.6780	0.684	0.6810	98.75
ImageCLEF	0.6802	0.775	0.7245	132.09
MS-COCO	0.7170	0.724	0.7205	217.65

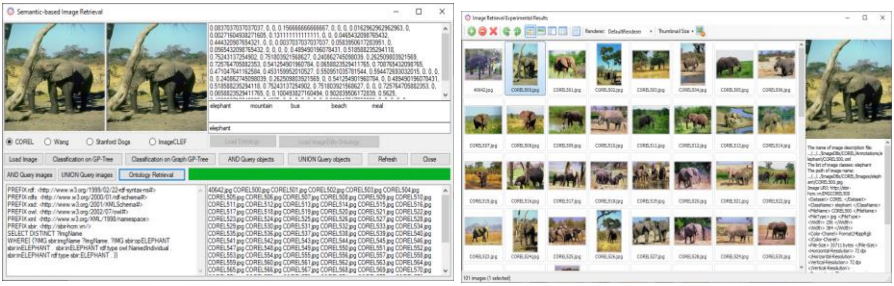

Figure 2.3. Image search system based on GP-Tree structure and search results

2.3. Chapter conclusion

This chapter describes the GP-Tree clustering tree structure, an effective solution for storing and retrieving large data, especially in image search. GP-Tree uses a hierarchical clustering method, which helps to quickly search through branches with high similarity. At the leaf nodes, the system identifies the most similar elements, optimizing time and achieving good accuracy. However, GP-Tree also has limitations, mainly when splitting nodes, similar elements can be distributed to different nodes, leading to reduced search performance, especially when they are no longer in the same branch in the tree. The next chapter will introduce improvements to GP-Tree to improve image search accuracy, focusing on optimizing the tree structure, reducing the omission of similar elements. At the same time, an image search model is also proposed to improve search performance.

CHAPTER 3. IMPROVING GP-TREE WITH GRAPH-GPTREE FOR IMAGE RETRIEVAL

This chapter presents a method to improve GP-Tree by combining it with the neighboring cluster graph, forming a Graph-GPTree structure. This improvement aims to improve the efficiency of storing and searching similar elements. An image search model based on Graph-GPTree is proposed, with the goal of improving the accuracy and speed of searching. Experiments are performed on the Wang, MS-COCO, and ImageCLEF datasets to evaluate the effectiveness of the method. Experiments on standard datasets such as Wang, MS-COCO, and ImageCLEF have demonstrated the effectiveness of the proposed method. The research results of this chapter are published in [CT2], [CT3], and are also further demonstrated through the content of [CT5] and [CT6].

3.1. The cluster graph Graph-GPTree

3.1.1. Graph-GPTree structure

Graph-GPTree is created based on operations on the sparse graph of the leaf nodes obtained from GP-Tree. In which, the vertices represent the leaf nodes and the weighted edges represent the similarity between them. The sparse graph is created during the process of creating a GP-Tree when each time the leaf node is split, the system marks the neighbor levels of the newly split leaf nodes (**Figure 3.1**)

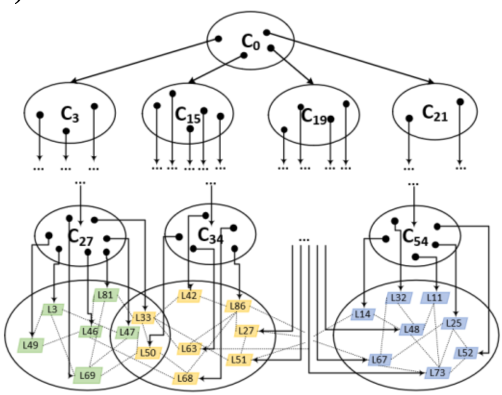

Figure 3.1. The sparse graph is created from the leaf node set of GP-Tree

3.1.2. Graph-GPTree construction process

The GraphGP-Tree graph clustering algorithm to find clusters in a data set is described in general as in **Figure 3.2.** The algorithm is performed on a sparse graph in which the nodes represent data elements and the weighted edges represent the similarity between data elements. Representing the data set using this sparse graph allows the clustering algorithm to scale to large data sets.

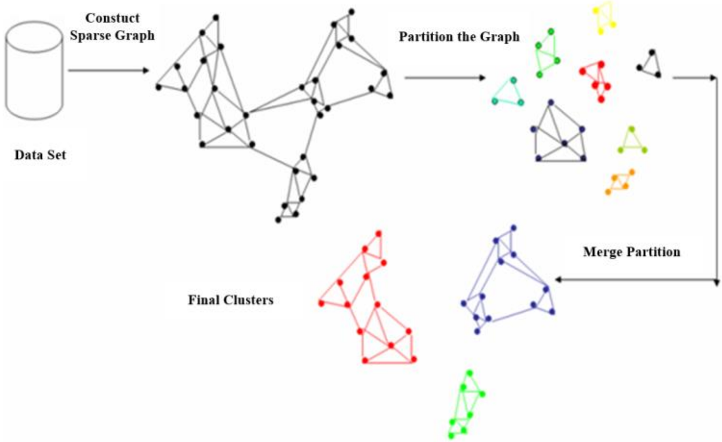

Figure 3.2. Overview of the overall approach of the GP-Tree algorithm

3.2. Image search system on Graph-GPTree

3.2.1. Experimental model

The image search model on the Graph-GPTree cluster graph (**Figure 3.3**) includes two main stages: the preprocessing stage and the image search stage, each stage plays an important role in optimizing the search results and improving the accuracy of the system. Preprocessing stage: (1) segment and extract features from the image data set; (2) build the Graph-GPTree model. Image search stage: (1) extract features from the input image; (2) compare the feature vector with the database on the GP-Tree to select similar branches and determine the appropriate leaf node; (3) From the leaf node, use the Graph-GPTree graph to find the neighborhood set and arrange similar images according to the similarity measure; (4) Apply k-NN algorithm to classify similar images and use visual vocabulary to interpret semantics.

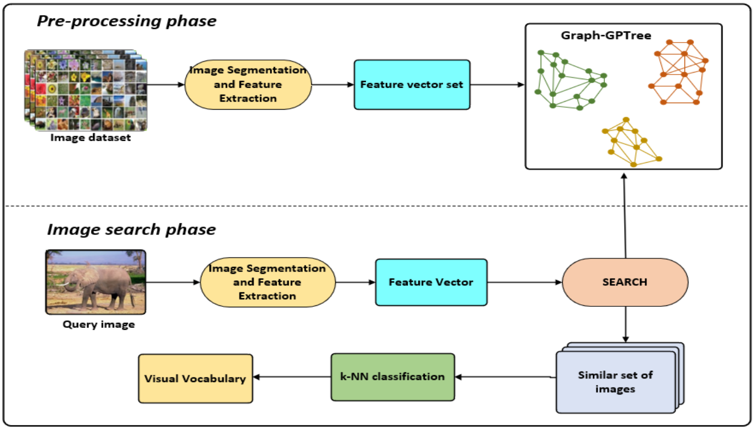

Figure 3.3. Graph-GPTree-based image retrieval model

3.2.2. Experiment and evaluation of SBIR-GP image retrieval system

SBIR-GP image retrieval system is designed to perform semantic-based image queries using SgGP-Tree and ontology. When given an input image, SBIR-GP system first extracts feature vectors and retrieves similar images based on their content sequentially using SgGP-Tree. This process results in a collection of similar images.

The image datasets used for the experiments include WANG, MS-COCO, and ImageCLEF datasets. The average performance values and search times of the test datasets are presented in **Table 3.1, Table 3.2** and **Table 3.3**

Table 3.1. Image search performance on WANG dataset

Method	Precision	Recall	F-Measure	Average Search Time (ms)
GP-Tree	0.6780	0.6840	0.6810	39.75
Graph-GPTree	0.7665	0.6677	0.7137	202.79

Table 3.2. Image search performance on ImageCLEF dataset

Method	Precision	Recall	F-Measure	Average Search Time (ms)
GP-Tree	0.6802	0.7750	0.7245	44.09
Graph-GPTree	0.8168	0.7637	0.7894	239.29

Table 3.3. Image search performance on MS-COCO dataset

Method	Precision	Recall	F-Measure	Average Search Time (ms)
GP-Tree	0.717	0.724	0.7205	102.32
Graph-GPTree	0.873	0.764	0.815	198.47

From the above tables, it can be seen that the improvement of GP-Tree helps to improve the exact search performance for the WANG, ImageCLEF and MS-COCO datasets. The Graph-GPTree neighborhood graph has better performance than GP-Tree. However, the search time of GP-Tree is faster than Graph-GPTree.

To evaluate the accuracy and efficiency of the Graph-GPTree-based image retrieval system, the performance of the system is compared with other works on the same dataset. The results in **Table 3.4** compare the proposed method with other research works on the WANG dataset. Graph-GPTree shows better accuracy than other methods, but still lower than the results of O. Sikha and K. Soman, and A. Ouni. Although the results have not improved much, the Graph-GPTree-based image retrieval methods still have quite good efficiency on the WANG image dataset. **Table 3.5** and **Table 3.6** compare the average accuracy of the search methods on the WANG, ImageCLEF and MS-COCO datasets.

Table 3.4. Comparison of Image Retrieval Methods on the WANG Dataset

Method	Average Precision
K. Kanwal và cộng sự, 2020	0.5067
H. Zeng và cộng sự, 2021	0.6600
O. Sikha và K. Soman, 2021	0.8030
S. Dhingra và P. Bansal, 2021	0.6000
A. Ouni và cộng sự, 2022	0.7800
Graph-GPTree	**0.7665**

Table 3.5. Comparison of Image Retrieval Methods on the ImageCLEF Dataset

Method	Average Precision
K. Kanwal và cộng sự, 2020	0.5067
H. Zeng và cộng sự, 2021	0.6600
O. Sikha và K. Soman, 2021	0.8030
S. Dhingra và P. Bansal, 2021	0.6000
A. Ouni và cộng sự, 2022	0.7800
Graph-GPTree	**0.7665**

Table 3.6. Comparison of Image Retrieval Methods on the MS-COCO Dataset

Method	Average Precision
Y. Cao và cộng sự, 2018	0.8576
Wen Gu và cộng sự, 2019	0.8350
Y. Xie và cộng sự, 2020	0.8628
Graph-GPTree	**0.8730**

The above tables show that the proposed method achieves higher accuracy than other retrieval methods, demonstrating the ability to extract features effectively and distinguish objects in detail in images. This method is effective in solving query and semantic analysis problems with both single and multi-object images.

3.3. Chapter Summary

This chapter proposes methods to improve image retrieval performance on GP-Tree. First, the Graph-GPTree model combines the neighborhood graph and GP-Tree, overcoming the limitation when similar elements are scattered through different branches during node separation. The use of the neighborhood graph helps to connect similar elements, enhancing the ability to search more accurately and efficiently.

CHAPTER 4. SGGP-TREE STRUCTURE FOR SEMANTIC IMAGE RETRIEVAL

This chapter proposes a new improvement to the GP-Tree structure by combining Graph-GPTree and self-organizing network (SOM), called SgGP-Tree. This structure is designed to improve the efficiency of image storage and retrieval. At the same time, a semantic image retrieval model based on ontology is proposed, called SBIR-GP. Experiments on standard datasets such as Wang, MS-COCO, and ImageCLEF have confirmed the superior performance of the method. The research results of this chapter are published in [CT3], and are also further demonstrated through the contents of [CT5] and [CT6].

4.1. SgGP-Tree *structure*

The SgGP-Tree structure is a combination of GP-Tree, Graph-GPTree and SOM network. In the SOM network, adjusting the weights during training will make SOM achieve the best clustering. The combined model of GP-Tree, Graph-GPTree neighbor graph and grSOM network, called SgGP-Tree, is described as in **Figure 4.1.**

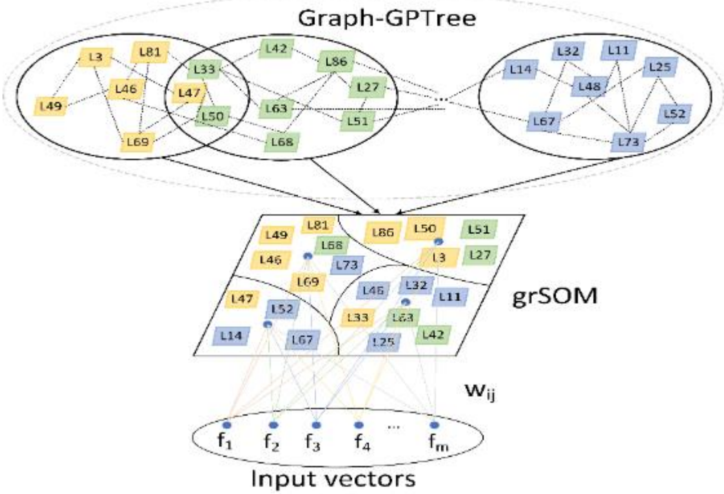

Figure 4.1. SgGP-Tree combined model

4.2. Image retrieval system on SgGP-Tree

4.2.1. Experimental model

The image retrieval model on the SgGP-Tree combined network is a combination of GP-Tree, Graph-GPTree neighbor cluster graph and SOM network (**Figure 4.2**). The preprocessing process includes extracting features from images and storing them on SgGP-Tree. During the search process, SgGP-Tree is used to find similar images and visual vocabulary. This model reuses the feature extraction block from previous models, with the improvement of combining Graph-GPTree graph and SOM network into GP-Tree, forming SgGP-Tree.

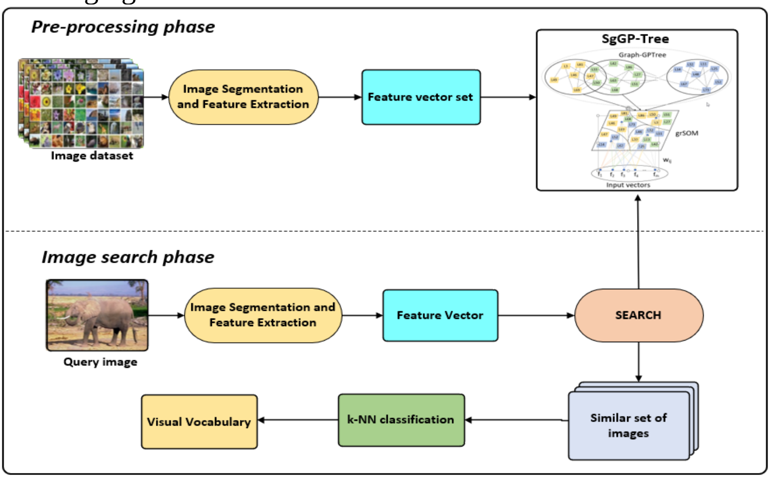

Figure 4.2. Image search model on SgGP-Tree

4.2.2. Experiment and evaluation of results

When receiving input images, the system extracts feature vectors and searches for similar images based on content using SgGP-Tree. The result is a set of images similar to the input image. **Figure 4.3** illustrates the interface of the image search system based on SgGP-Tree with input images and describes the results of the similar image set retrieved from the searched image.

The image datasets used for the experiments include WANG, MS-COCO and ImageCLEF datasets. The average performance values and search times of the test dataset are presented in **Table 4.1, Table 4.2** and **Table 4.3**

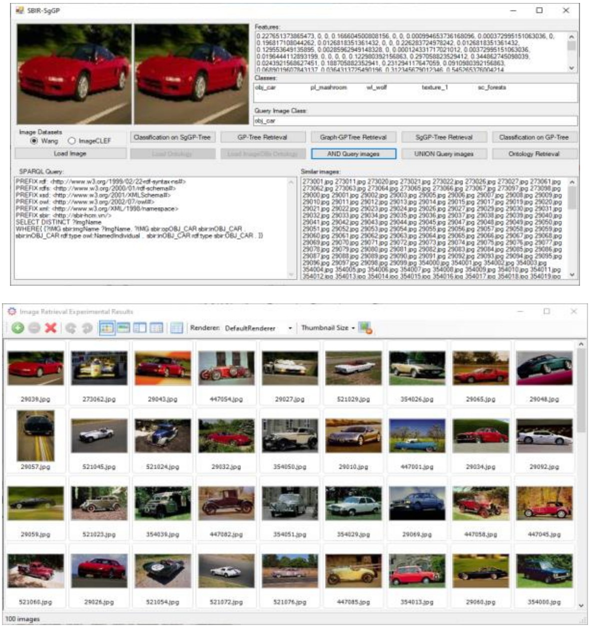

Figure 4.3. Interface of the SgGP-Tree semantic image retrieval system

Table 4.1. Image search performance on WANG dataset

Method	Precision	Recall	F-Measure	Average Search Time (ms)
GP-Tree	0.6780	0.6840	0.6810	39.75
Graph-GPTree	0.7665	0.6677	0.7137	202.79
SgGP-Tree	0.8004	0.7040	0.7491	696.19

Table 4.2. Image search performance on ImageCLEF dataset

Method	Precision	Recall	F-Measure	Average Search Time (ms)
GP-Tree	0.6802	0.7750	0.7245	44.09
Graph-GPTree	0.8168	0.7637	0.7894	239.29
SgGP-Tree	0.8926	0.8764	0.8844	868.51

Table 4.3. Image search performance on MS-COCO dataset

Method	Precision	Recall	F-Measure	Average Search Time (ms)
GP-Tree	0.717	0.724	0.7205	102.32
Graph-GPTree	0.873	0.764	0.815	198.47
SgGP-Tree	0.875	0.724	0.783	265.45

From the above tables, it can be seen that the improvement of GP-Tree improves the exact search performance for the WANG, ImageCLEF and MS-COCO datasets. The neighborhood graph Graph-GPTree has a better performance than GP-Tree but lower than SgGP-Tree. However, the search time of GP-Tree is faster than Graph-GPTree and SgGP-Tree.

To evaluate the performance of the search system, the ROC and PR curves are used. The area under the curve (AUC) in the ROC space is a measure of the accuracy of the search process, with the larger the area, the higher the accuracy. The PR curve combines precision and coverage, with AUC similar to ROC. The ROC and PR curves are used to evaluate the accuracy of the SgGP-Tree search system, as shown in **Figure 4.4, Figure 4.5** and **Figure 4.6**

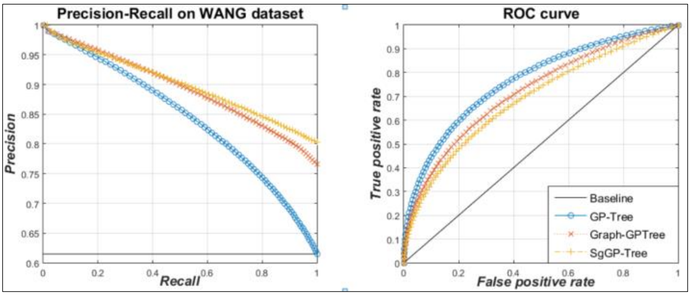

Figure 4.4. Image search performance on GP-Tree, Graph-GPTree and SgGP-Tree on the WANG image dataset.

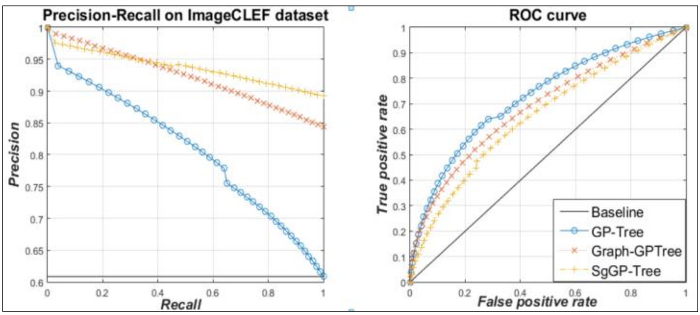

Figure 4.5. Image search performance on GP-Tree, Graph-GPTree and SgGP-Tree on the ImageCLEF image dataset.

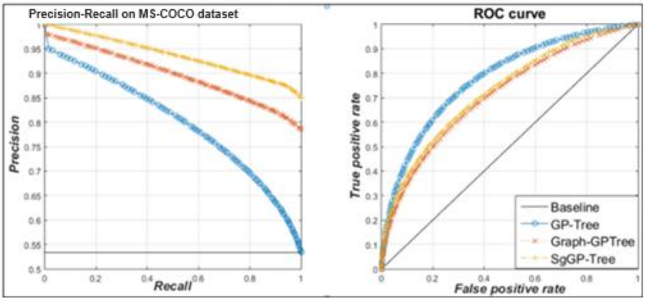

Figure 4.6. Image search performance on GP-Tree, Graph-GPTree and SgGP-Tree on the MS-COCO image dataset.

To evaluate the accuracy and efficiency of SgGP-Tree, the performance of the system is compared with other works on the same dataset. The results in **Table 4.4, Table 4.5** and **Table 4.6** compare the average accuracy of the search methods on the WANG, ImageCLEF and MS-COCO datasets.

Table 3.4. Comparison of Image Retrieval Methods on the WANG Dataset

Method	Average Precision
H. Zeng và cộng sự, 2021	0.6600
O. Sikha và K. Soman, 2021	0.8030
S. Dhingra và P. Bansal, 2021	0.6000
A. Ouni và cộng sự, 2022	0.7800
SgGP-Tree	**0.8004**

Table 3.5. Comparison of Image Retrieval Methods on the ImageCLEF Dataset

Method	Average Precision
Y. Qiang và cộng sự, 2020	0.6670
X. Yue và cộng sự, 2021	0.7140
N. T. U. Nhi và cộng sự, 2022	0.6510
X. Wang và cộng sự, 2023	0.7270
SgGP-Tree	**0.8926**

Table 3.6. Comparison of Image Retrieval Methods on the MS-COCO Dataset

Method	Average Precision
Y. Cao và cộng sự, 2018	0.8576
Y. Xie và cộng sự, 2020	0.8628
Wen Gu và cộng sự, 2019	0.8350
SgGP-Tree	**0.8753**

The above tables show that the proposed method achieves higher accuracy than other retrieval methods, demonstrating the ability to effectively extract features and distinguish details of objects in images. This method is effective in solving query and semantic analysis problems with both single and multi-object images.

4.3. Semantic image retrieval system based on ontology

4.3.1. Ontology framework

The ontology is built and developed based on the set of object images and the relationships between objects. The initial multi-object images are segmented into object images; the components of the object images are extracted and the relationships between objects are built. **Figure 4.7** illustrates the ontology built on Protégé for MS-COCO data

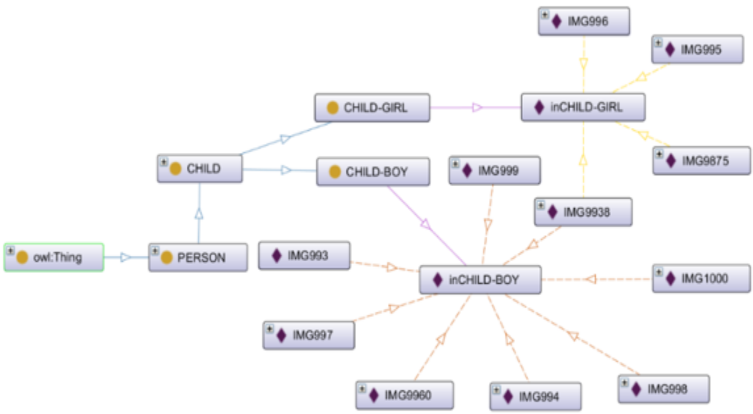

Figure 4.7. An example of ontology applied to the MS-COCO image dataset

4.3.2. Semantic-based image retrieval model

SBIR-GP semantic image retrieval system is a combination of SgGP-Tree machine learning structure and ontology, aiming to improve the ability to search and classify images based on semantics (**Figure 4.8**). This system consists of two stages: pre-processing stage and image retrieval stage. Each stage plays an important role in optimizing search results and improving the accuracy of the system.

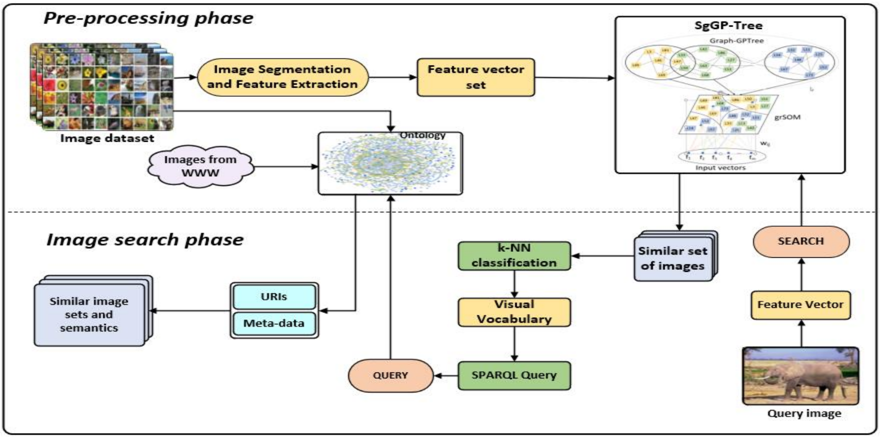

Figure 4.8. SBIR-GP semantic image retrieval model

4.3.3. Experiment and evaluation of results

SgGP-Tree semantic image retrieval system uses SgGP-Tree and ontology to search for similar images based on semantics. When receiving input images, the system extracts feature vectors and searches for similar images based on content using SgGP-Tree. The image datasets used for the experiments include the WANG, MS-COCO, and ImageCLEF datasets. The average performance values and search times of the test datasets are shown in **Table 4.7**

Table 4.7. Image search performance of the SBIR-GP system on the test datasets

Image Dataset	**Evaluation Metrics**			
	Accuracy	**Recall**	**F-measure**	**Average Search Time (ms)**
WANG	0.884	0.754	0.814	214,72
ImageCLEF	0.943	0.837	0.887	276,33
MS-COCO	0.915	0.841	0.877	312,65

From **Figure 4.9**, it can be seen that the GP-Tree enhancement improves the exact search performance for the WANG, ImageCLEF, and MS-COCO datasets. The SBIR-GP system outperforms GP-Tree, Graph-GPTree, and SgGP-Tree, demonstrating its ability to efficiently extract features and distinguish detailed objects in images. This method is effective in solving query and semantic analysis problems with both single and multi-object images.

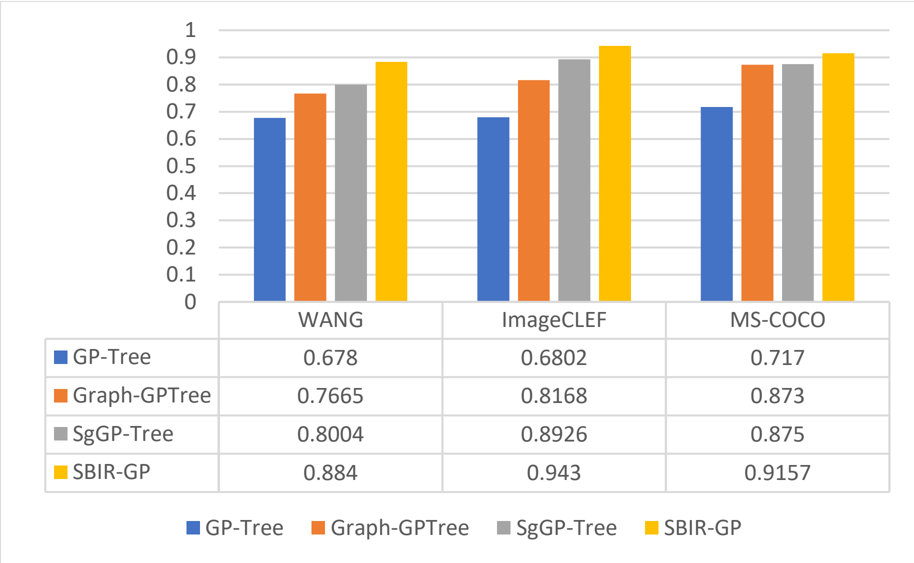

Figure 4.9. Comparison of image search performance on WANG, ImageCLEF and MS-COCO image datasets

4.4. Chapter Summary

In this chapter, methods to improve the image retrieval performance on GP-Tree are proposed. First, a model combining the neighbor graph with GP-Tree, called Graph-GPTree, is created to connect similar elements that are branched during node splitting on GP-Tree. Next, a model combining GrSOM and Graph-GPTree, called SgGP-Tree, is created to improve the image retrieval performance. The SgGP-Tree model adds criteria to select winning leaf nodes, which helps to cluster better and search for images more accurately. Experiments are conducted on the WANG, ImageCLEF, and MS-COCO datasets. The SBIR-GP system has superior accuracy compared to previous proposals by the author of the thesis. The experimental 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 SBIR-GP retrieval system has higher accuracy than other studies on the same image dataset. This proves that the proposals in the thesis are effective and appropriate.

CONCLUSION AND DEVELOPMENT DIRECTIONS

❖ Contributions of the thesis

The thesis proposes and develops semantic-based image retrieval methods with 4 main contributions:

- GP-Tree structure: building a GP-Tree structure based on hierarchical clustering, storing and indexing image data effectively, reducing data size, increasing search speed and accuracy.
- Graph-GPTree structure: Developing Graph-GPTree combining neighborhood graphs, improving storage and retrieval efficiency, solving the data dispersion phenomenon in GP-Tree.
- SgGP-Tree structure: Combining Graph-GPTree with SOM network, optimizing image storage and querying, reducing the complexity of large data processing and improving query speed.
- SBIR-GP model: Proposing a semantic image search method based on ontology and SgGP-Tree, overcoming the limitations of traditional methods, improving efficiency when processing complex context images.

These contributions have been verified on standard datasets and published in scientific works, affirming the potential for application in fields such as medicine, tourism, and social networks.

❖ Development direction

Although important results have been achieved, the thesis still has some development directions that can be expanded and improved the efficiency of the semantic image search system:

- Comparison with modern methods: Evaluate the effectiveness of deep learning models such as CNN, R-CNN, and GCN to improve search accuracy and performance.
- Practical applications: Expand into areas such as tourism, medicine, soil, and social networks to increase practical value.
- Enriching Knowledge Graph (KG): Strengthening semantic relationships in KG to improve image search capabilities in complex cases.
- Developing Vietnamese KG: Building KG suitable for Vietnamese semantics, supporting domestic applications and meeting the needs of the Vietnamese market.

LIST OF PUBLISHED ARTICLES RELATED TO THE THESIS

1. **N. M. Hai**, T. V. Lang, and V. T. Thanh, "Semantic-Based Image Retrieval Using Hierarchical Clustering and Neighbor Graph," in World Conference on Information Systems and Technologies, 2022, pp. 34-44: Springer, DOI: https://doi.org/10.1007/978-3-031-04829-6_4 (**Scopus, Q4**)
2. **N. M. Hai**, V. T. Thanh, and T. V. Lang, "A method for semantic-based image retrieval using hierarchical clustering tree and graph," Telkomnika, vol. 20, no. 5, pp. 1026-1033, 2022, DOI: http://doi.org/10.12928/telkomnika.v20i5.24086 (**Scopus, Q3**)
3. **N. M. Hai**, T. Van Lang and T. T. Van, "Improving The Efficiency of Semantic Image Retrieval using A Combined Graph and SOM Model," in IEEE Access, 2023, doi: https://doi.org/10.1109/ACCESS.2023.3333678 (**SCIE, Q1**)
4. **N. M. Hai**, V. T. Thanh, and T. V. Lang, "The improvements of semantic-based image retrieval using hierarchical clustering tree," FAIR'2020, 2020, pp. 557-570: Natural Science and Technology Publishing House, DOI: https://doi.org/10.15625/vap.2020.00213
5. **Nguyễn Minh Hải**, Trần Văn Lăng, Văn Thế Thành, "Một tiếp cận tìm kiếm ảnh theo ngữ nghĩa dựa trên mạng nơ-ron tích chập và ontology", TCKHTrường ĐH Sư phạm TP. HCM, 2022. tr. 48-59. DOI: https://doi.org/10.54607/hcmue.js.19.3.3272(2022)
6. **N. M. Hai**, V. T. Thanh, and T. V. Lang, "A method of semantic-based image retrieval using graph cut," Computer Science and Cybernetics, vol. 38, no. 2, pp. 193-212, 2022, DOI: https://doi.org/10.15625/1813-9663/38/2/16786