

Research Article

Ship-SIBISCaS: A First Step towards the Identification of Potential Maritime Law Infringements by means of LSA-Based Image

Gabriel Alejandro León-Paredes ¹, Liliana Ibeth Barbosa-Santillán ²,
Juan Jaime Sánchez-Escobar ³ and Antonio Pareja-Lora ⁴

¹Universidad Politécnica Salesiana, Cuenca, Ecuador

²Universidad de Guadalajara, Jalisco, Mexico

³Centro de Enseanza Técnica Industrial, Jalisco, Mexico

⁴Universidad Complutense de Madrid, Madrid, Spain

Correspondence should be addressed to Juan Jaime Sánchez-Escobar; jjsanchez@ceti.mx

Received 8 August 2018; Accepted 27 January 2019; Published 3 March 2019

Academic Editor: Giuseppe Scanniello

Copyright © 2019 Gabriel Alejandro León-Paredes et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Maritime safety and security are being constantly jeopardized. Therefore, identifying maritime flow irregularities (semi-)automatically may be crucial to ensure maritime security in the future. This paper presents a Ship Semantic Information-Based, Image Similarity Calculation System (Ship-SIBISCaS), which constitutes a first step towards the automatic identification of this kind of maritime irregularities. In particular, the main goal of Ship-SIBISCaS is to automatically identify the type of ship depicted in a given image (such as abandoned, cargo, container, hospital, passenger, pirate, submersible, three-decker, or warship) and, thus, classify it accordingly. This classification is achieved in Ship-SIBISCaS by finding out the similarity of the ship image and/or description with other ship images and descriptions included in its knowledge base. This similarity is calculated by means of an LSA algorithm implementation that is run on a parallel architecture consisting of CPUs and GPUs (i.e., a heterogeneous system). This implementation of the LSA algorithm has been trained with a collection of texts, extracted from Wikipedia, that associate some semantic information to ImageNet ship images. Thanks to its parallel architecture, the indexing process of this image retrieval system has been accelerated 10 times (for the 1261 ships included in ImageNet). The range of the precision of the image similarity method is 46% to 92% with 100% recall (that is, a 100% coverage of the database).

1. Introduction

Nowadays, many maritime law infringements dealing with vessels and cargo (or trading) ships involve (i) the irregular transit of this kind of ships through areas where their presence is absolutely restricted or, alternatively, and (ii) their berthing, mooring, loading, and/or unloading in areas not conceived for this purpose. For example, (a) the transit of cruises and other recreational crafts can lead to severe environmental damage, and thus, it is forbidden in protected sea environments (e.g., the surroundings of Cozumel island); and (b) the transit, berthing, mooring, loading, and/or

unloading of vessels and/or trading and cargo ships in sea recreational areas is usually associated with shipping black market transactions. Therefore, a number of maritime security procedures have to be implemented, in order to prevent this kind of irregularities from happening whenever possible and, when not possible, to immediately detect and stop them.

So far, current maritime security procedures have relied mainly on humans (i.e., coast guards). However, maritime law infringements can be pervasive, the areas to control are immense, and suitable human and transportation resources to fight them are usually scarce. Thus, human-based maritime

security procedures are being proven insufficient in order to take these irregularities under control. For this reason, there is currently an increasing need for the deployment of some additional (semi-)automatic procedures that can supplement them and eventually ensure maritime protection at least in the most problematic areas.

One of the additional procedures that can and may be developed towards detecting ship transit irregularities is to analyze (semi-)automatically some images taken by one or more satellites of certain maritime zones. In fact, it is globally accepted that “satellite (...) observation services are indispensable in any effective, modern maritime monitoring system. Satellite image data are essential to ensure continuous maritime monitoring” [1]. In this case, the main goal of image analysis is to detect when a given kind of ship is in an irregular (or restricted) area. This main goal can be decomposed in three different subgoals or problems, namely, (1) identifying the kind of ship in question, that is, classifying the ship; (2) identifying the type of area in which it is located; and (3) checking if the ship is transiting an allowed or forbidden area for its kind.

This paper presents Ship-SIBISCaS, a prototype system that has served as a proof of concept and as a first attempt towards solving subgoal (1), i.e., towards classifying a ship located in an area of interest. Ship-SIBISCaS is a hybrid text and image retrieval system that allows classifying a ship by means of two different types of inputs, that is, (a) a text defining and/or describing it or (b) an image depicting it. The outputs of the system are both the class to which the ship belongs and the ships in its image database that are most similar to (or, ideally, match) the input text or image associated to it. This retrieval system comprises both a software and a hardware subsystem. On one hand, the Ship-SIBISCaS software subsystem has been built around an instance of the Latent Semantic Analysis (LSA) algorithm, particularly trained for this purpose. On the other hand, its hardware subsystem is a heterogeneous, parallel hardware architecture (i.e., consisting of both CPUs and GPUs) previously developed [2], but conveniently customized to suit this new use case.

As shown by the results of the experiments carried out with the Ship-SIBISCaS system, (1) applying the LSA algorithm to solve this problem helped avoiding the disadvantages of some previous approaches to image matching and/or retrieval, while attaining a satisfactory level of precision and recall, whereas (2) reusing the heterogeneous, parallel hardware architecture allowed speeding-up the execution of the LSA algorithm and achieving a (nearly) real-time response of the system when it is queried.

The rest of the paper describes mainly the construction of the system and the experiments carried out in order to evaluate it. It has been organized as follows. Firstly, Section 2 provides an overview of the related works. Secondly, Section 3 presents how the Ship-SIBISCaS system was designed in order to retrieve ship images, commenting also its main design issues and requirements. Thirdly, Section 4 describes the most relevant details of its development and the experiments performed to test and evaluate it. Fourthly, Section 5 shows the results of these experiments. Finally, Section 6 summarizes the conclusions drawn from this work.

2. Related Works

In the last years, more and more ship and/or vessel data have been generated, shared, and made available online. These ship data are represented as text, image, audio, or video in different formats. In particular, the number of private or public ship digital image collections is rapidly increasing. These image data and collections are currently being used, e.g., in the field of visual information processing, to provide robust models for many types of image processing, such as ship recognition and classification.

A neural network is a classifier that needs a previous training to predict the class to which an object (an image, in this case) belongs. The main disadvantage of neural networks is that their training often fails due to the existence of a local optimum [3]. Nevertheless, when they can be trained, they often get good results as for precision. In any case, the neural network approach did not fulfill this task requirements. In effect, in addition to training the model for classification, it was required that the algorithm added to images their semantic description. This allowed enriching their associated information and enabled further and enhanced postprocessing. Therefore, neural networks had to be discarded for this research.

Smeulders et al. [4] propose new directions for researchers in this particular field. One of them is how to identify the image features (i) that suitably represent the meaning of an image; (ii) that can label the image to describe it appropriately; and (iii) that can be used to retrieve the image when necessary. The research field dealing with this kind of problems is known as Content-Based Image Retrieval (CBIR). One of the main problems that CBIR researchers have to face is what they refer to as the *semantic gap*. In this context, the *semantic gap* can be defined as “the lack of coincidence between the information that one can extract from the visual data and the interpretation” that a particular user makes “of the same data (...) in a given situation.”

To reduce this semantic gap, several works have been carried out so far [5–8]. As shown by these works, there are two main types of features that can be used to characterize and retrieve images, namely, visual and textual features. On the one hand, images can be retrieved by means of their visual features, such as color, shape, or texture [9–11]. On the other hand, images can be retrieved by means of their textual features (that is, features of the text describing them) such as keywords, tags, captions, and annotations [12–14].

Besides, other works in this area have already shown that using only one type of feature may result in poor outcomes. For instance, some authors report that using exclusively visual features does not help describe the full semantics of an image [15]. In addition, some other authors have proven that CBIR systems based just on textual features or on visual features often yield irrelevant results [16, 17] since the quality of image textual descriptions depends on (a) the skills of the people who write them (for instance, their knowledge, intelligence, experience, and lexicon) and (b) the different potential interpretations that an image may have.

Fortunately, the Latent Semantic Analysis (LSA) method has successfully overcome the problem of text retrieval, for

texts with a conceptual content [18]. Furthermore, the LSA method has also been used to retrieve images with quite promising results [17, 19, 20, 21]. However, the computational complexity and the memory requirements of the LSA method have been shown to be extremely high over the years [22–24], especially when its Singular Value Decomposition (SVD) process is considered. Besides, image similarity assessment and/or image matching are high time-consuming tasks too, due to the high computational cost of their associated indexing and retrieval processes. Thus far, this extremely high complexity has made it almost impossible to apply LSA to image retrieval over large image collections, due to memory restrictions. Therefore, it would be most useful to provide a solution to this complexity problem, thus enabling the execution of LSA over image collections of any size.

Some of the most recent works in the areas mentioned above deserve some special attention, due to their relevance with respect to the present work. They are the following:

- (i) Anandh et al. proposes a technique for the generation of image content descriptors with the features of a color autocorrelogram, Gabor wavelet, and wavelet transform in order to reduce the semantic gap in CBIR. They achieved accuracy rates of 83% for the Corel database, 88% for the Li database, and 70% for the Caltech-101 database [9].
- (ii) Han et al. describes a semantic image processing mechanism for automatic context-awareness based on cloud computing. They applied some powerful high computation resources to represent computationally image semantics and/or how image similarities are perceived. A semantic inference is done through user-created multimedia contents that are added to the images. User's perception and user semantically defined images are matched against each other and, as a consequence, images can be classified according to user's definitions and their semantics [12].
- (iii) Wan et al. shows the development of a framework of deep learning for CBIR tasks. More specifically, some experiments were carried out in this work to assess the suitability of state-of-the-art deep convolutional neural networks for CBIR tasks and thus find out if deep learning has the potential to bridge the semantic gap in CBIR. For object image retrieval, the authors evaluated a dataset on the Caltech256 image database and obtained results that demonstrated that some pretrained models were able to capture some highly semantic information from raw pixels [25].
- (iv) Stathopoulos and Kalamboukis presents an innovative application of the LSA algorithm that can skip the need to calculate the full SVD of the feature matrix and, hence, that overcomes the deficiencies of this algorithm when run on large-scale datasets. This application was tested with a collection provided at the ImageCLEFs 2012 & 2013 Medical Image Retrieval congresses. The

results showed that visual techniques alone are not capable of fulfilling the semantic information needs of users [17].

- (v) Jing et al. provides an approach to the development of cost-effective and large-scale visual search systems using distributed computation platforms and open-source tools. In one of their experiments, the authors used as input the image annotations generated in the Pinterest scheme, such as pin descriptions and board titles. These image annotations provide a great deal of text information about the image. Hence, the experiment showed that the task of object detection using both text and visual data achieved a very low rate of false positives (less than 1%) [16].

In particular, the present work is most similar to [12] because it also uses (i) semantics and classification according to definitions (though Wikipedia and WordNet were used in our case instead) and also (ii) collaborative tagging of the ship images, in order to enrich their semantics. However, this work combines LSA and parallel processing in order to accelerate the computation of the processes involved. As with [17], this helps overcome the deficiencies of applying the LSA algorithm on large-scale datasets, but in a more robust and general way.

3. Design Issues and Requirements of the Ship Semantic Information-Based, Image Similarity Calculation System (Ship-SIBISCaS)

As previously introduced, this paper presents a Ship Semantic Information-Based, Image Similarity Calculation System (Ship-SIBISCaS) that applies the LSA algorithm to image retrieval and provides a solution to overcome the complexity problem of this algorithm when run over large image collections. The main objective of Ship-SIBISCaS is, thus, to retrieve the ships in a collection related to a specific kind of ship in order to (i) calculate their similarity and (ii) enable the classification of a particular query ship with the aid of the similar ships retrieved. The results obtained with this first prototype of Ship-SIBISCaS will be applied in the future to try and ensure maritime security (semi-) automatically.

The development of Ship-SIBISCaS has had to face two main challenges, namely, (1) the generation of its knowledge base with a convenient image dataset as input and its enrichment by linking the images to some appropriate text and/or semantic features, and (2) learning semantic information about these images from the text and/or features linked to them. An overview of the processes used in the Ship-SIBISCaS is shown in Figure 1.

3.1. Generating and Enriching the Knowledge Base: Image Dataset Acquisition and Text and/or Semantic Feature Linking. One of the most important requirements for training a machine-learning algorithm is to use accurate data

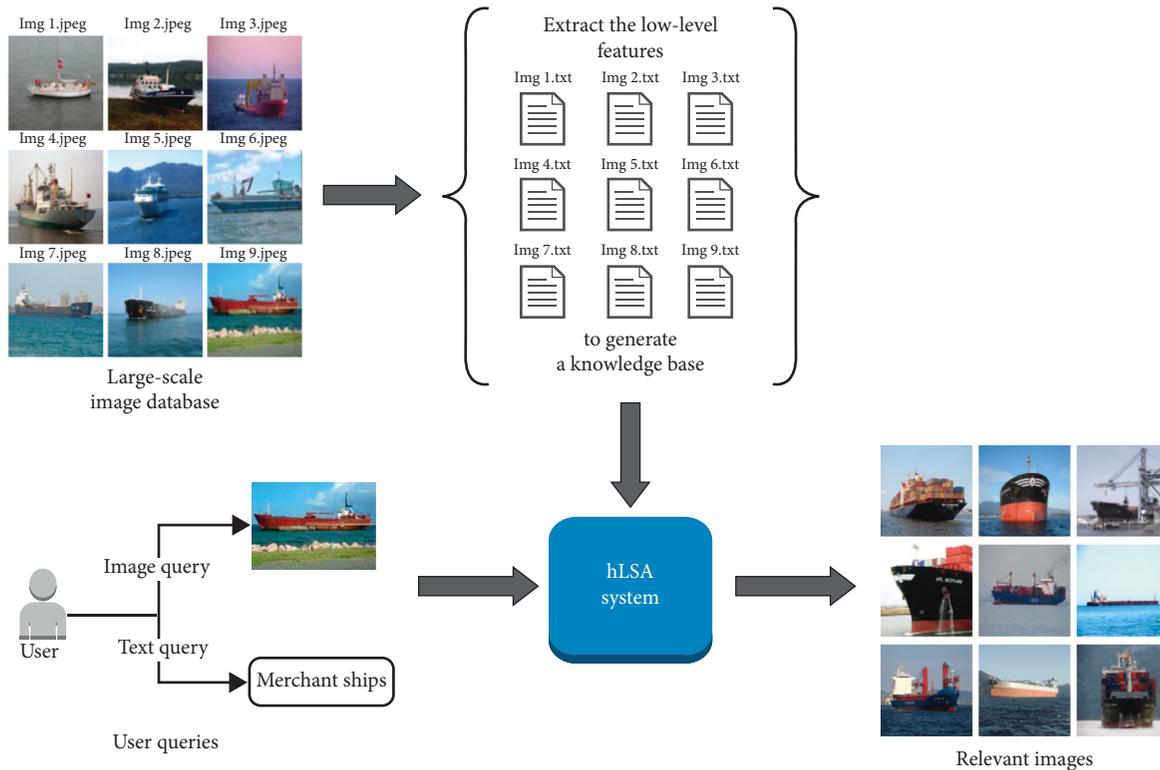


FIGURE 1: An overview of the processes implemented in Ship-SIMI.

as input. Indeed, if the machine-learning algorithm processes ambiguous or incorrect images when being trained, then it will, for instance, associate incorrect textual information to images. For this reason, when the Ship-SIBISCaS system was designed it was decided (a) to populate its knowledge base with a ship image dataset conveniently cleaned in order to reduce the noise and (b) to have this ship image dataset enriched by associating textual features to the ship images. It was devised also that this should be performed by means of the following four stages: (1) image dataset crawling, (2) collaborative tagging, (3) linking, and (4) cleaning.

3.1.1. Image Dataset Crawling. The input dataset for the population of the knowledge base had to be ImageNet [26]. In ImageNet, each node of the hierarchy is composed of hundreds or thousands of images that could allow Ship-SIBISCaS image dataset to be built. Towards this aim, only the ImageNet nodes corresponding to ships would have to be selected and incorporated into the knowledge base of Ship-SIBISCaS. The knowledge base would have to be enriched afterwards by extracting from WordNet (WordNet®, <https://wordnet.princeton.edu>) the lexical features associated to its images.

3.1.2. Collaborative Tagging. An online collaborative tagging system (<http://cloudcomputing.ups.edu.ec/ImageTagProject>) would have to be built in order to attach some annotations to the images in the database. By using this tagging system, a

human annotator could associate a ship image with one or more tags, which should describe as faithfully as possible the type of ship depicted in the image. This information would be saved in a structured database within the knowledge base containing the identifier of the image (ID), its name, its URL, its category, and a tag for the image, among other information related to the ship collaborative tagging system.

3.1.3. Linking. The information stored in the structured database would then be used to obtain further tagging associations between an image and one or more tags. Then, a web search engine would have to be developed and applied to crawl the World Wide Web in order to attach some textual information to the ship image-tags relations. This textual information, in general, would consist of definitions or descriptions of the type of ship depicted in the image (as detailed in the image tag).

3.1.4. Cleaning. The textual information gathered by the web search engine would then be automatically cleaned in order to remove all unnecessary noisy paragraphs and characters and to obtain thus a clean knowledge base for the Ship-SIBISCaS system. Hence, the knowledge base would consist of text documents, where the file name of each document should correspond to the file name of the image, and its image-tags relations.

This knowledge base would be used afterwards to train a semantic space within the Ship-SIBISCaS system. Figure 2 shows how Ship-SIBISCaS can generate and enrich the

knowledge base with the textual information associated to the ship image dataset.

3.2. How to Learn Semantic Information with the Ship-SIBISCaS LSA Algorithm. As already mentioned in Section 1, to train Ship-SIBISCaS and retrieve ship images from its knowledge base, an existing heterogeneous latent semantic analysis (hLSA) system [2] had to be reused to implement a data-intensive instance of the LSA algorithm.

This hLSA system uses heterogeneous architectures (that is, including both CPUs and GPUs) in order to accelerate the execution of the LSA method, especially to enable it to run on large-scale datasets. It has been developed using (i) GPU computing to solve faster large numerical problems by means of the thousands of concurrent threads on the multiple CUDA core multiprocessors and (ii) multi-CPU computing to solve faster text problems via their shared memory-programming model in a multiprocessing environment.

The hLSA system can train a knowledge base and retrieve information in less than two minutes for a collection of 5,000 documents or, equivalently, a term-document matrix containing one hundred and fifty thousand million values. Thus, the hLSA system has been used in Ship-SIBISCaS to train the knowledge base and to retrieve similar ship images in real-time. Table 1 shows the acceleration attained by the hLSA system when compared to the classical LSA system using 5000 input documents, two weighting schemes (namely, Log Entropy and TF-IDF), and a reduction to $k = 300$ dimensions for the three use cases presented in [2].

The hLSA system is divided into three main stages, as shown in Figure 3. The first stage creates the semantic space, the second stage reduces conveniently its dimensionality (that is, the k value), and the third stage retrieves the requested information from the semantic space associated to the database.

3.2.1. Semantic Space Creation. The hLSA system preprocesses the text in order to remove any remaining strange characters, blank spaces, and common words (stopwords) from the text document knowledge base. Then, the hLSA system generates a term-document matrix where each entry indicates the frequency of a given term in a given document. The generation of this term-document matrix has a high computational cost due to the large amounts of text included in the knowledge base. Thus, the hLSA system implements a parallel data-intensive algorithm that shares the memory in the multi-CPU architecture. This is done in order to reduce the processing time required to index and construct the term-document matrix.

3.2.2. Dimensionality Reduction. The resulting term-document matrix is fairly large, and it is populated mostly with zeros. Therefore, it needs to be, respectively, reduced and normalized. To normalize it, the hLSA system utilizes well-known term-weighting metrics, such as the logarithmic local number (Log Entropy) and the term

frequency-inverse document frequency (TF-IDF) formula. To reduce the normalized matrix, the hLSA system uses instead a truncated Singular Value Decomposition (SVD) of the matrix to reflect its major associative patterns in a smaller number of dimensions (characterized by the k value, that is, the number of resulting rows or columns in the reduced matrix). The SVD algorithm has a high computational cost for large matrices, and therefore the hLSA system implements a parallel data-intensive SVD algorithm using the benefits of GPU architectures to accelerate its calculation.

3.2.3. Information Retrieval. The two stages above are used to learn to calculate similarity measures and classify images from the knowledge base in order to obtain a reduced semantic space. Then, the similarity values are ranked in a descending order, where the most similar documents have the highest values and the least relevant documents have the lowest values. Therefore, a supervised learning model is developed in order to retrieve ship images that are relevant for a user query. The hLSA system compares each document vector in the term-document matrix with the user query vector. Additionally, the hLSA system presents the relevant images and documents (texts) in a graphic user interface as shown in Figure 4.

4. Details of Development of Ship-SIBISCaS and Experiments Carried Out So Far

This section describes in detail, firstly, the most relevant issues in the development of the main design components of the Ship-SIBISCaS system (described in Section 3) and, secondly, the experiments carried out so far in order to test and evaluate it.

To start with, the most relevant details in the development of Ship-SIBISCaS are discussed below, namely, (1) how the ship image dataset was acquired, (2) how the tagset to annotate the ship image was determined, (3) how the ship images were collaboratively tagged, (4) how the knowledge base was enriched by linking the images to suitable texts, (5) how the semantic space associated to the knowledge base was trained, and (6) how similar images are retrieved from the trained semantic space.

- (1) *Image dataset acquisition.* It was decided to use the ImageNet database as main input [26] to test and evaluate this first prototype of Ship-SIBISCaS. ImageNet consists of nodes that have been organized following the structure provided by WordNet noun synsets and their corresponding hierarchical relationships. Thus, in this database, each node contains hundreds or thousands of images whose content can be described by the representative noun of a WordNet synset. Therefore, the database has more than 100,000 nodes and has an approximately 14 million images. ImageNet provides around 1,000 images per synset that are quality-controlled and human-annotated. In order to create and generate the knowledge base of Ship-SIBISCaS, only the ship

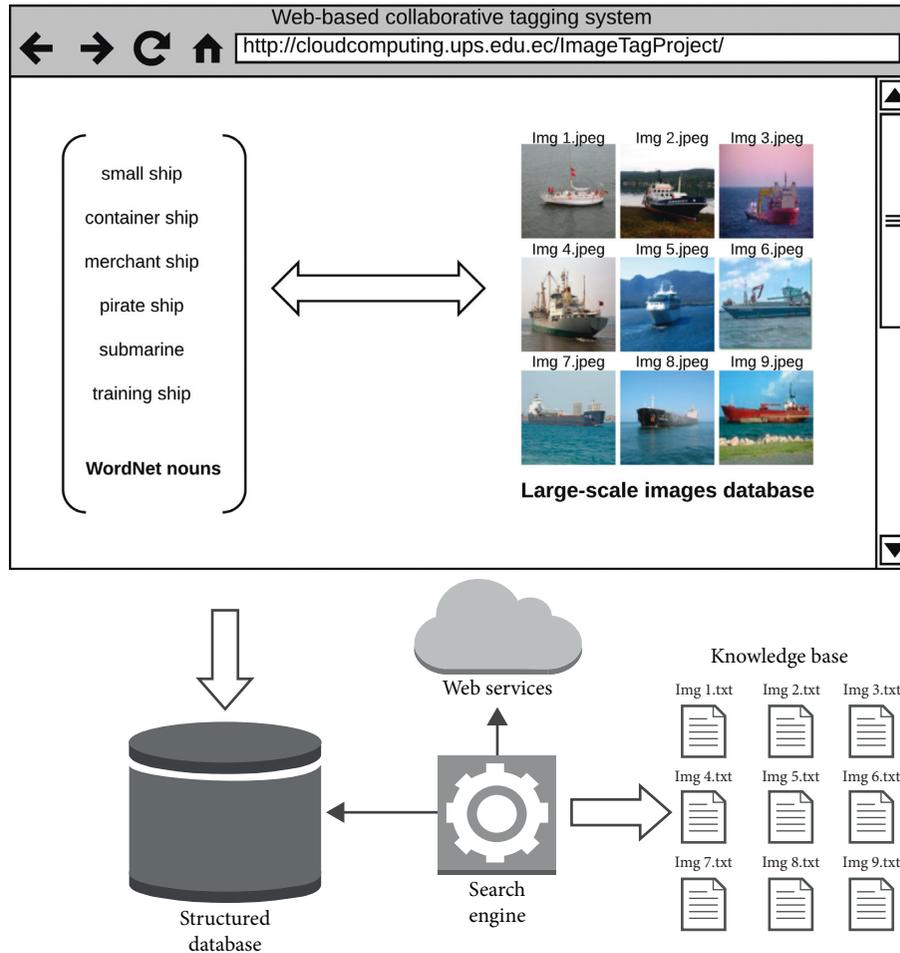


FIGURE 2: The collaborative tagging process used to enrich the knowledge base of Ship-SIBISCaS.

TABLE 1: A comparison of the execution times of a classical LSA system vs. the hLSA system using 5000 input documents, two weighting schemes, and $k=300$ dimensions for three use cases.

| | Bipolar disorders | | Lupus disease | | Topiramate weight loss | |
|--------------|-------------------|--------------|-------------------|--------------|------------------------|--------------|
| | Log entropy (sec) | TF-IDF (sec) | Log entropy (sec) | TF-IDF (sec) | Log entropy (sec) | TF-IDF (sec) |
| hLSA | 96.29 | 110.16 | 87.42 | 86.40 | 181.93 | 180.25 |
| LSA | 770.50 | 610.82 | 756.33 | 590.01 | 848.97 | 689.12 |
| Acceleration | 8.00 | 5.54 | 8.65 | 6.83 | 4.67 | 3.82 |

collection of images was extracted from ImageNet. This collection is defined as “a vessel that carries passengers or freight,” and contains 78 child synsets, such as abandoned ship, cargo ship, cargo vessel, hospital ship, passenger ship, and transport ship, among others. The ship collection comprises a total of 1,261 images. Figure 5 shows a part of the taxonomy and some images related to this ship image collection.

- (2) *Tagset determination.* Still, the number of nouns that could be used to tag (that is, to classify) ship images was far too large. In order to reduce the number of the tags to be eventually used for this purpose, an

online collaborative tagging system was developed. The main aim of this system was to determine which are the most frequently used and/or most useful WordNet nouns for describing these ship images. Then, this collaborative system was used by a set of human annotators (referred to as a focus group here), and it was found out that only 43 terms were required to tag and classify the ships depicted in the ImageNet ship image collection. These 43 tags (the Ship-SIBISCaS tagset) are shown in Table 2.

- (3) *Collaborative tagging.* Then, the online collaborative tagging system was retargeted to annotate the images in the Ship-SIBISCaS knowledge base using the

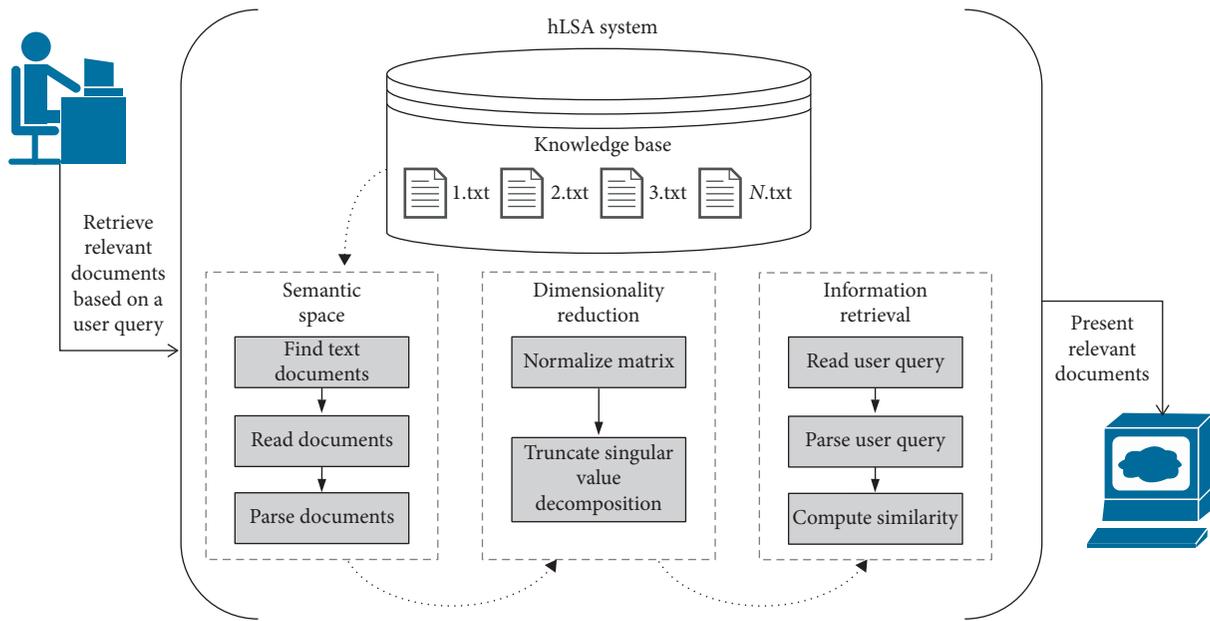


FIGURE 3: An overview of the three stages within the hLSA system.

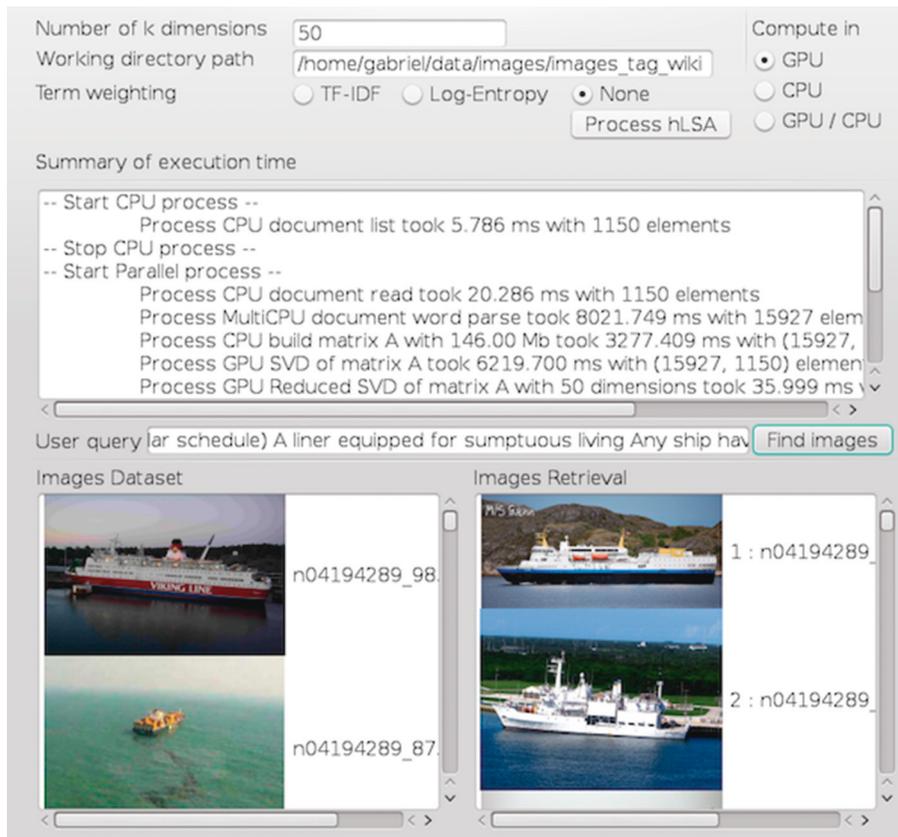


FIGURE 4: The hLSA system presents the relevant images and documents in a graphic user interface.

tagset previously determined. For this purpose, a controlled focus group of approximately 120 people with intermediate knowledge in the field of ship images was hired. Each person in the group tagged an average of 20 images. A total of 2,320 associations

between one image and its tag were obtained this way. The following considerations must be made: (a) the same image could be tagged by different people, so a given image can have several (repeated) associations; and (b) the members of the focus group

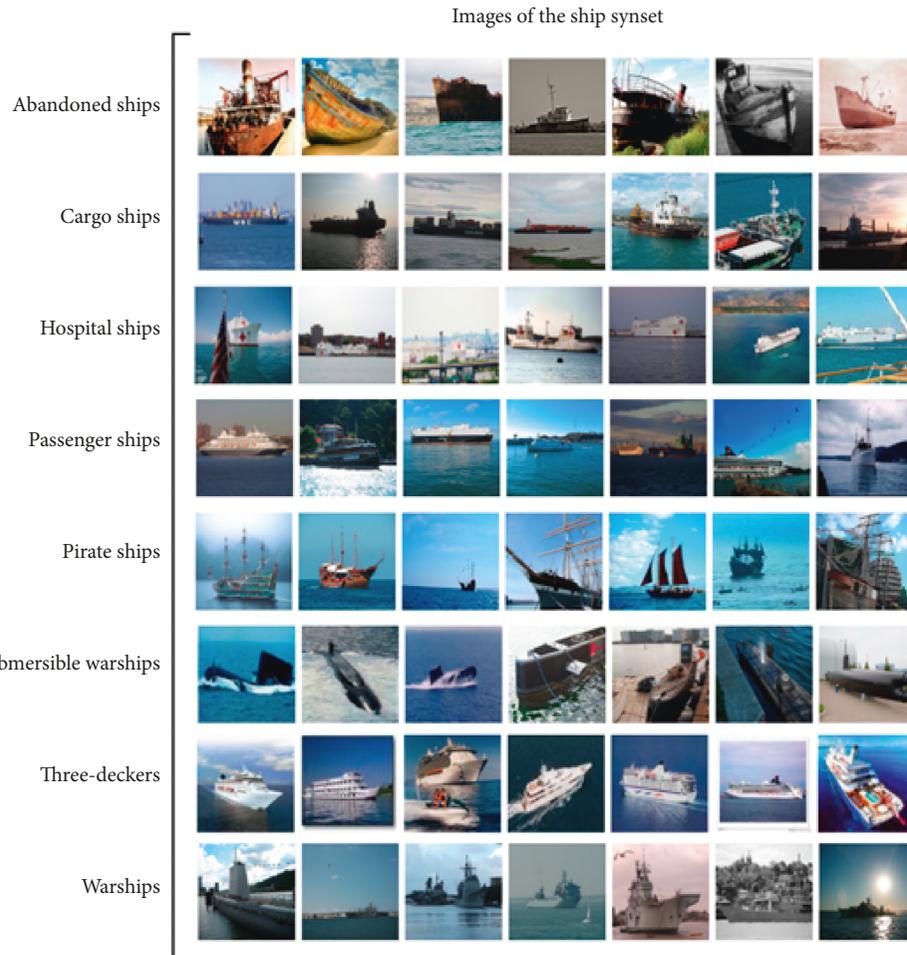


FIGURE 5: A part of the taxonomy and some images of the ImageNet “ship” related collection.

tagged the images according to their knowledge of the field and, therefore, their corresponding associations (tags) for an image can coincide or not. As a result, at least 711 images were associated with a single tag and at least 550 images were associated with more than one tag.

- (4) *Linking.* The Ship-SIBISCaS system was used then to link the image by means of its corresponding tags to some suitable text descriptions. This was achieved by means of its web-based search engine, which uses a set of web services to find the textual information to be linked to an image via its associated tag(s). In particular, this engine was used to obtain textual information from Wikipedia (Wikipedia (<http://wikipedia.com>) is a well-known online free encyclopedia) through its API (Application Programming Interface). Also, the web-based search engine enriched the knowledge base with further semantic information, such as descriptions. The file name of each document in the knowledge base corresponds to the name of the image to which it is linked, and it contains some relevant textual information for each image-tag association.

- (5) *Semantic space (LSA) training.* The knowledge base, enriched with the image tags and the image textual information and/or descriptions, was then used as input to train the hLSA system described in the previous section. As commented, the hLSA system reduces the dimensionality of the matrix A (or, equivalently, of the semantic space) by using a truncated SVD of A in order to reflect its major associative patterns and ignore the smaller, less important influences. It restricts the dimensionality of matrix A to its first k dimensions, where k must be smaller than the minimum among the total number of terms and documents, that is, $k < \min(\text{terms}, \text{documents})$. The number of dimensions k was equal to 50. The execution time obtained for this process was 8 seconds.
- (6) *Retrieving similar images.* The system uses the cosine measure in order to obtain the similarity between an image vector of the trained semantic space and a query vector. Ship-SIBISCaS retrieves similar images for an image query or for a text query. For image queries, it uses an indexed image of the trained semantic space as a query, and for the text query, it uses the tag descriptions presented in Table 2. The

TABLE 2: The Ship-SIBISCaS tagset for the annotation and/or classification of ImageNet ship images.

| | Tag name | Tag description |
|----|------------------------|--|
| 1 | Abandoned ship | A ship abandoned on the high seas |
| 2 | Aircraft carrier | A large warship that carries planes and has a long flat deck for takeoffs and landings |
| 3 | Banana boat | A ship designed to transport bananas |
| 4 | Battle cruiser | A cruiser of maximum speed and firepower |
| 5 | Cargo ship | A ship designed to carry cargo |
| 6 | Container ship | A cargo ship designed to hold containerized cargoes; "the weight of the documentation of all the consignments on board a contemporary container ship can exceed 90 pounds" |
| 7 | Cruise ship | A passenger ship used commercially for pleasure |
| 8 | Factory ship | A whaling ship equipped to process whale products at sea |
| 9 | Fishing vessel | A fishing vessel is a boat or ship used to catch fish in the sea, or on a lake or river. Many different kinds of vessels are used in commercial and recreational fishing. |
| 10 | Flagship | The ship that carries the commander of a fleet and flies his flag |
| 11 | Guided missile frigate | A frigate that carries guided missiles |
| 12 | Hospital ship | A ship built to serve as a hospital; used for wounded in wartime |
| 13 | Icebreaker | A ship with a reinforced bow to break up ice and keep channels open for navigation |
| 14 | Ironclad | A wooden warship of the 19th century that is plated with iron or steel armor |
| 15 | Lightship | A ship equipped like a lighthouse and anchored where a permanent lighthouse would be impractical |
| 16 | Luxury liner | A liner equipped for sumptuous living |
| 17 | Merchant ship | A cargo ship; "they did much of their overseas trade in foreign bottoms" |
| 18 | Minesweeper | Ship equipped to detect and then destroy or neutralize or remove marine mines |
| 19 | Ocean liner | A large commercial ship (especially one that carries passengers on a regular schedule) |
| 20 | Oil tanker | A cargo ship designed to carry crude oil in bulk |
| 21 | Passenger ship | A ship built to carry passengers |
| 22 | Patrol boat | A vessel assigned to patrol an area. |
| 23 | Pirate ship | A ship that is manned by pirates |
| 24 | Privateer | A privately owned warship commissioned to prey on the commercial shipping or warships of an enemy nation |
| 25 | Research submarine | A submarine for research purposes |
| 26 | Sailing boat | A sailboat or sailing boat is a boat propelled partly or entirely by sails smaller than a sailing ship |
| 27 | Sailing ship | A modern sailing ship is any large wind-powered vessel. Traditionally a sailing ship (or simply ship) is a sailing vessel that carries three or more masts with square sails |
| 28 | Sailing warship | A warship that was powered by sails and equipped with many heavy guns; not built after the middle of the 19th century |
| 29 | Shipwreck | A wrecked ship (or a part of one) |
| 30 | Slave ship | A ship used to transport slaves from their homes to places of bondage |
| 31 | Small ship | A ship that is small |
| 32 | Steamship | A ship powered by one or more steam engines |
| 33 | Submarine | A submersible warship usually armed with torpedoes |
| 34 | Supply ship | A ship that usually provides supplies to other ships |
| 35 | Surface ship | A warship that operates on the surface of the water |

TABLE 2: Continued.

| | Tag name | Tag description |
|----|---------------|---|
| 36 | Three-decker | Any ship having three decks |
| 37 | Torpedo boat | Small high-speed warship designed for torpedo attacks in coastal waters |
| 38 | Training ship | A ship used to train students as sailors |
| 39 | Treasure ship | A 16th century ship loaded with treasure |
| 40 | Troopship | A ship for carrying soldiers or military equipment |
| 41 | Warship | A government ship that is available for waging war |
| 42 | Wooden boat | A wooden boat operating near shore |
| 43 | Yacht | A medium-sized sailboat equipped for cruising or racing |

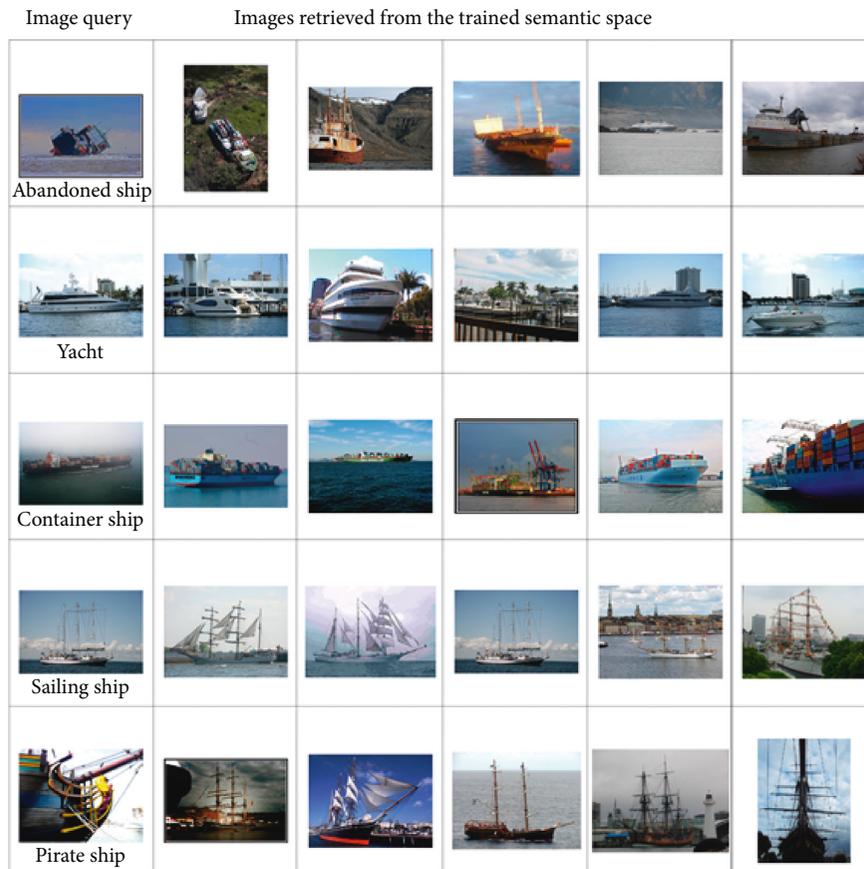


FIGURE 6: Image queries: most similar images retrieved from the semantic space for five relevant test cases.

method compares all the vectors of the trained semantic space with the query vector. As a result, the five images with the highest similarity values are retrieved. The execution time obtained for this process varies depending on the semantic information of the query. Execution times of 10 to 30 seconds were obtained in the case of image queries and of 100 to 300 milliseconds in the case of the text queries. Figure 6 shows five image queries and their most similar images in the trained semantic space. Besides, Figure 7 shows five text queries and the most similar images retrieved for them from the semantic space.

5. Results

First, we present the results from collaborative tagging process, followed by the size metrics associated to the ImageNet sample used in our experiments. Then, we show an estimation of the performance of Ship-SIBISCaS by means of the usual precision and recall measures. Finally, we detail the results of the system concerning its execution times.

Hence, firstly, the results concerning the collaborative tagging process for each ship tag (or category) are shown in Figure 8.

As shown in Table 3, we found a total of 15,921 unique words in the knowledge base. Therefore, the index process of

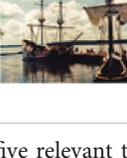
| Text query | Images retrieved from the trained semantic space | | | | |
|---|--|--|--|---|--|
| <p>Abandoned ship</p> <p>A ship abandoned on the high seas</p> |  |  |  |  |  |
| <p>Yacht</p> <p>A medium-sized sailboat equipped for cruising or racing</p> |  |  |  |  |  |
| <p>Container ship</p> <p>A cargo ship designed to hold containerized cargoes ...</p> |  |  |  |  |  |
| <p>Sailing ship</p> <p>A modern sailing ship is any large wind-powered vessel ...</p> |  |  |  |  |  |
| <p>Pirate ship</p> <p>A ship that is manned by pirates</p> |  |  |  |  |  |

FIGURE 7: Text queries: most similar images retrieved from the semantic space for five relevant test cases.

the knowledge base generated a term-document matrix A with 15,921 terms. This index process was performed for 1,261 documents. It should be remarked that the matrix A was populated mostly by zeros. Thus, we normalized the matrix A by using the term-weighting scheme known as Term Frequency-Inverse Document Frequency (TF-IDF). The final total number of values in the matrix A is 20,076,381, which corresponds to 161 megabytes. This process was executed in 9.5 seconds.

Secondly, the size metrics associated to the ImageNet sample used in our experiments are shown in Table 3, where the number of unique words is 15,921.

Thirdly, we show an estimation of the performance of Ship-SIBISCaS by means of the usual precision and recall measures. Basically, precision values give information about the effectiveness of the system, whereas recall values provide some information as for the coverage of the system, as shown in Table 4. For example, the text query for “yacht” returned some false-positives because the associated tag description was incorrect, with a precision of 2%. However, the image query process uses a better description for container ships, and their associated answers get an average precision of 87%.

Lastly, the results concerning the execution times associated to the matrix reduction process and the retrieval of ship images based on semantic information are shown in Figure 9.

To conclude the description of the experiments, it must be said that the Ship-SIBISCaS software subsystem has been tested on a GPU configuration with a global memory of 3064 MBytes and a clock speed of 2.5 GHz. Also, the maximum number of threads per multiprocessor is 2048 and the maximum number of threads per block is 1028. The GPU maximum dimension size of a thread block is (1024, 1024, 64) in the (x, y, z) dimensions.

6. Conclusions

This paper has introduced a Ship Semantic Information-Based, Image Similarity Calculation System (Ship-SIBISCaS) that helps classifying a ship depicted in an image. This classification is regarded here as a proof of concept and a first step towards the (semi-)automatic identification of maritime flow irregularities and, thus, as a potentially most helpful mechanism to ensure maritime security.

To achieve this goal, firstly, the ship-related image dataset of the ImageNet database was extracted and loaded into the knowledge base around which the system has been built. This ship image dataset comprised 1,261 images of different ship categories.

Secondly, the ship images in this dataset have been annotated by means of an online collaborative tagging system developed on purpose. This has helped make explicit 2,320 associations between ship images and annotations.

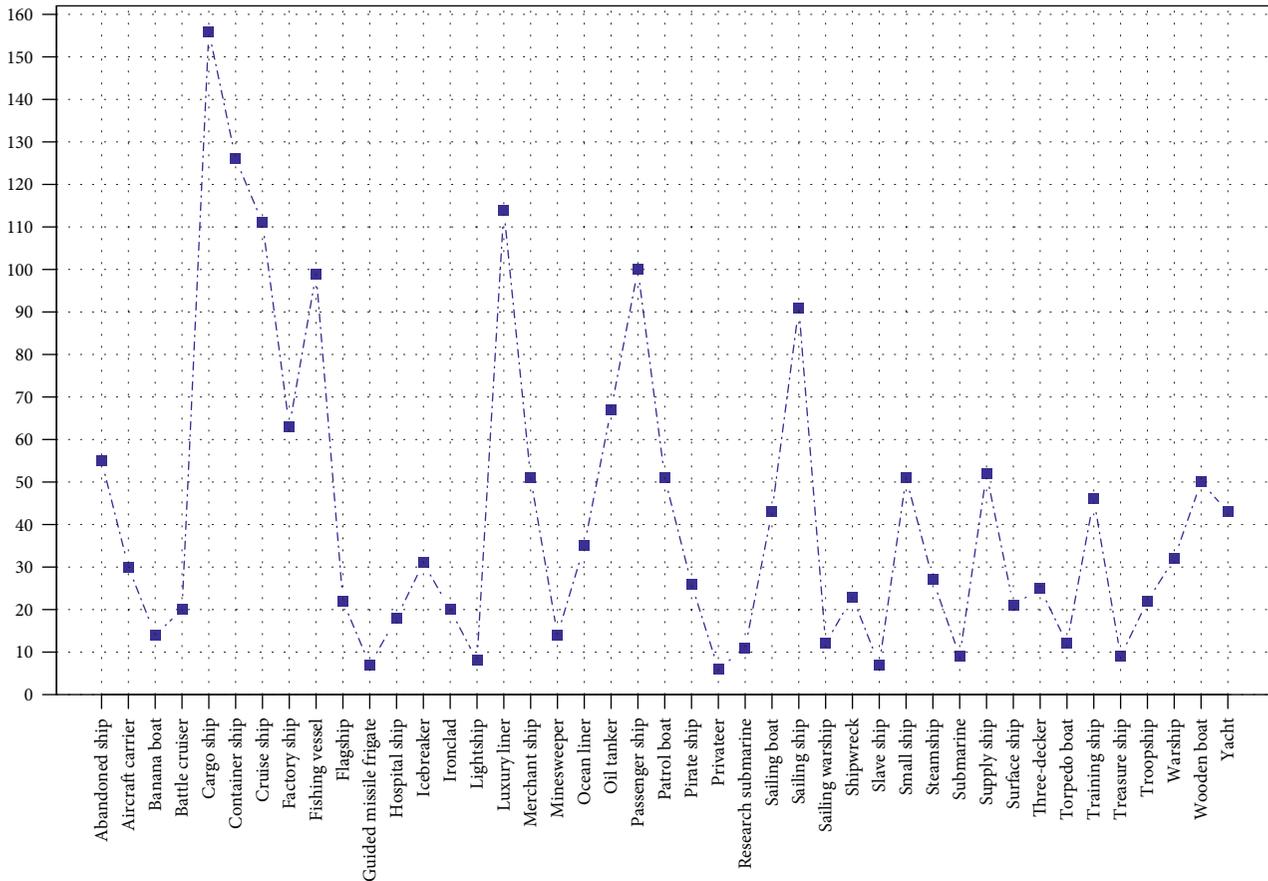


FIGURE 8: Collaborative tagging results for each of the ship tags.

TABLE 3: Sample size of the Ship-SIBISCaS system.

| Number of documents | Number of words | Number of unique words | Size of knowledge base (Mb) |
|---------------------|-----------------|------------------------|-----------------------------|
| 1,261 | 5,767,357 | 15,921 | 35 |

TABLE 4: Recall and precision results obtained in the experiments carried out with Ship-SIBISCaS.

| Tag | Image query | | Text query | |
|----------------|-------------|---------------|------------|---------------|
| | Recall (%) | Precision (%) | Recall (%) | Precision (%) |
| Abandoned ship | 100 | 73 | 100 | 73 |
| Yacht | 100 | 77 | 0 | 2 |
| Container ship | 100 | 92 | 100 | 83 |
| Sailing ship | 100 | 46 | 100 | 46 |
| Pirate ship | 100 | 92 | 100 | 92 |

Thirdly, the semantic reasoner of Ship-SIBISCaS has gathered some textual information related to the ship image-tag associations in order to enrich the knowledge base with text descriptions.

Fourthly, an accelerated data-intensive algorithm (hLSA) has been used to train the system by learning the semantics of these text descriptions associated to the ship images.

Finally, the trained system has been successfully applied in some experiments (i) to classify the input ship (which can either be depicted in an image, or be described by means of

some text); and (ii) to retrieve images of similar ships from the knowledge base. On the one hand, when the query consists of a ship image, the system can provide an answer in 10–30 seconds. On the other hand, when the query consists of some textual description of the ship, the system can answer in only 100–300 milliseconds, due to the different types of semantic information applied by the reasoner in each case. Indeed, for the image query, Ship-SIBISCaS uses an indexed image, whereas a semantic description is used for the text query.

These experiments have also helped prove the following:

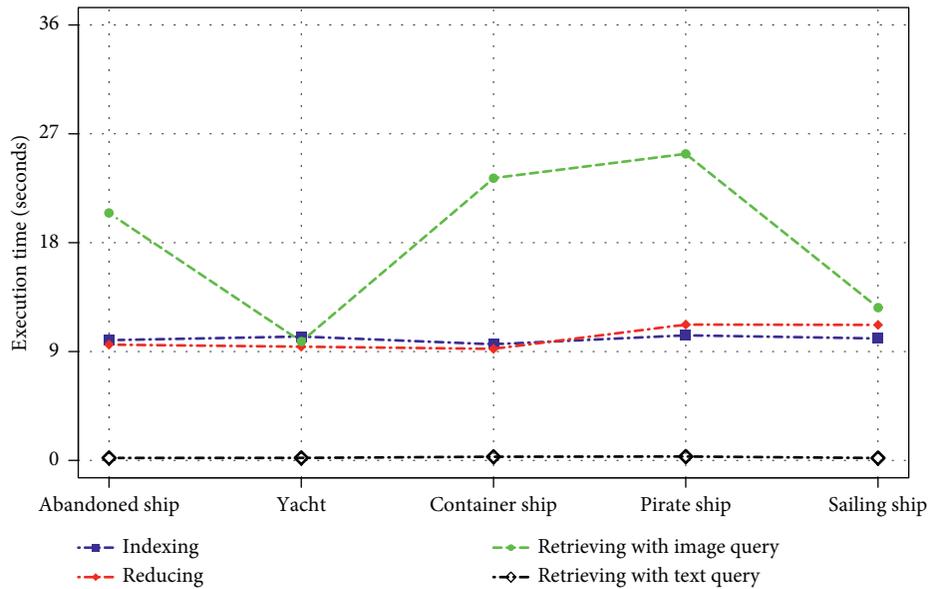


FIGURE 9: Amounts of time required to index, reduce, and retrieve similar ship images using a data-intensive LSA algorithm in Ship-SIBISCaS.

- (1) It is possible to automatically identify the type of ships, that is, if it is, e.g., an abandoned ship, a container ship, a cargo ship, a hospital ship, or a passenger ship.
- (2) Reusing within Ship-SIBISCaS the hLSA architecture that had already been developed to implement the LSA algorithm can accelerate the execution of this algorithm and make it run even 10 times faster.

Concerning the system performance, it took 17.5 seconds to train the model with the 1,261 input elements (ships). The resulting term-document matrix contains 20,076,381 values and has a size of 161 megabytes. Besides, more specifically, it took 9.5 seconds to execute the indexing phase and 8 seconds (averagely) to retrieve the answers (with $k = 50$ dimensions).

To conclude, as for the quality of the answers provided by the system, they may be described as satisfactory, but improvable. On the one hand, a recall of 100% has been obtained in the two experiments carried out with Ship-SIBISCaS. On the other hand, the Ship-SIBISCaS precision values have ranged from 2% to 92%. Most remarkably, the image similarity algorithm has had a better precision with image queries than with text queries. For instance, the container ship achieved 92% precision with image querying vs. 83% precision with text querying. However, Ship-SIBISCaS retrieves false-positives for text queries due to the noise associated to the image collaborative tagging process. In effect, some annotators annotated the same image differently (even wrongly), depending on their corresponding interpretation of the image, which was used later on to tag it and make its semantics explicit.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

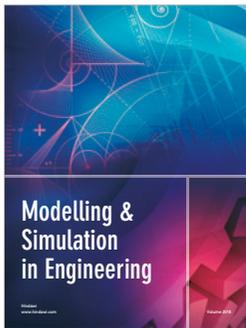
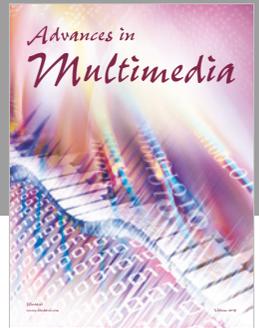
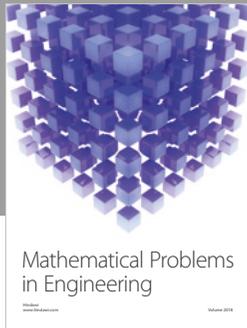
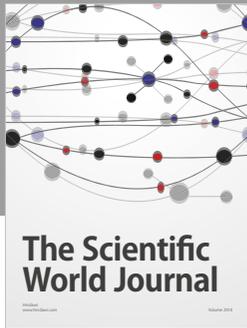
Acknowledgments

This work has been supported by the Sciences Research Council (CONACYT) through the research project no. 262756 called “The use of GNSS data for tracking maritime flow for sea security.” The authors thank their colleagues from the Universidad Politécnica Salesiana (UPS) through the research group Cloud Computing, Smart Cities & High Performance Computing (GIHP4C) and the project TIN2014-52010-R called RedR+Human by the Spanish Ministry of Economy and Competitiveness.

References

- [1] T. Stoltenberg, “Proposals presented to the extraordinary meeting of Nordic foreign ministers in Oslo on 9 February 2009,” Technical report, Finnish, Icelandic and Swedish Ministries of Foreign Affairs, Oslo, Norway, 2009, http://www.fride.org/download/OP_Stoltenberg_mante_paz_eng_jul09.pdf.
- [2] G. A. León-Paredes, L. I. Barbosa-Santillán, and J. J. Sánchez-Escobar, “A heterogeneous system based on latent semantic analysis (hlsa) using gpu and multi-cpu,” *Scientific Programming*, vol. 2017, Article ID 8131390, 19 pages, 2017.
- [3] O. Ludwig Jr. and U. Nunes, “Improving the generalization properties of neural networks: an application to vehicle detection,” in *Proceedings of 2008 11th International IEEE Conference on Intelligent Transportation Systems*, pp. 310–315, Beijing, China, October 2008.
- [4] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, “Content-based image retrieval at the end of the early

- years,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349–1380, 2000.
- [5] Y. H. Ali and W. N. Abdullah, “A survey of similarity measures in web image search,” *International Journal of Emerging Trends Technology in Computer Science*, vol. 4, no. 4, 2016.
- [6] A. Alzu’bi, A. Amira, and N. Ramzan, “Semantic content-based image retrieval: a comprehensive study,” *Journal of Visual Communication and Image Representation*, vol. 32, pp. 20–54, 2015.
- [7] X. Li, T. Uricchio, L. Ballan, M. Bertini, C. G. M. Snoek, and A. Del Bimbo, “Socializing the semantic gap: a comparative survey on image tag assignment, refinement, and retrieval,” *ACM Computing Surveys*, vol. 49, no. 1, 2016.
- [8] C. F. Tsai, “Bag-of-words representation in image annotation: a review,” *ISRN Artificial Intelligence*, vol. 2012, Article ID 376804, 19 pages, 2012.
- [9] A. Anandh, K. Mala, and S. Suganya, “Content based image retrieval system based on semantic information using color, texture and shape features,” in *Proceedings of 2016 International Conference on Computing Technologies and Intelligent Data Engineering, ICCTIDE*, Kovilpatti, India, January 2016.
- [10] C. W. Niblack, R. Barber, W. Equitz et al., “QBIC project: querying images by content, using color, texture, and shape,” in *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, vol. 1908, SPIE, Washington, DC, USA, April 1993.
- [11] J. Yue, Z. Li, L. Liu, and Z. Fu, “Content-based image retrieval using color and texture fused features,” *Mathematical and Computer Modelling*, vol. 54, no. 3-4, pp. 1121–1127, 2011.
- [12] S. H. Han, H. W. Kim, B. K. Park, Y. A. Heo, and Y. S. Jeong, “Efficient semantic image processing mechanism for automatic context-aware based on cloud infrastructure,” in *Advanced Multimedia and Ubiquitous Engineering*, Springer, Singapore, 2016.
- [13] S. Sclaroff, M. L. Cascia, S. Sethi, and L. Taycher, “Unifying textual and visual cues for content-based image retrieval on the World Wide web,” *Computer Vision and Image Understanding*, vol. 75, no. 1/2, pp. 86–98, 1999.
- [14] X. Zhou, “CBIR: from low-level features to high-level semantics,” in *Proceedings of SPIE—The International Society for Optical Engineering 2000*, University Press, New York, NY, USA, April 2000.
- [15] T. L. Berg, A. C. Berg, J. Edwards et al., “Names and faces in the news,” in *Proceedings of 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Washington, DC, USA, June–July 2004.
- [16] Y. Jing, D. Liu, D. Kislyuk et al., “Visual search at pinterest,” in *Proceedings of 21th ACM SIGKDD International Conference*, Sydney, NSW, Australia, August 2015.
- [17] S. Stathopoulos and T. Kalamboukis, “Applying latent semantic analysis to large-scale medical image databases,” *Computerized Medical Imaging and Graphics*, vol. 39, pp. 27–34, 2015.
- [18] S. Deerwester, S. Deerwester, S. T. Dumais et al., “Indexing by latent semantic analysis,” *Journal of the American Society for Information Science*, vol. 41, no. 6, pp. 391–407, 1990.
- [19] M. Bakalem, N. Benblidia, and S. Oukid, “Latent semantic analysis-based image auto annotation,” in *Proceedings of 2010 International Conference on Machine and Web Intelligence (ICMWI)*, Algiers, Algeria, October 2010.
- [20] C. H. Lee and K. C. Chiang, “Latent semantic analysis for classifying scene images,” in *Proceedings of International Multi Conference of Engineers and Computer Scientists 2010, IMECS*, Hong Kong, June 2010.
- [21] P. Zhang, Y. Zhang, T. Thomas, and S. Emmanuel, “Moving people tracking with detection by latent semantic analysis for visual surveillance applications,” *Multimedia Tools and Applications*, vol. 68, no. 3, pp. 991–1021, 2014.
- [22] S. T. Dumais, “Latent semantic analysis,” *Annual Review of Information Science and Technology*, vol. 38, no. 1, pp. 188–230, 2004.
- [23] J. R. Herrera-Morales and L. I. Barbosa-Santillán, “Analysis of medical publications with latent semantic analysis method,” in *Proceedings of Third International Conference on Advances in Information Mining and Management*, pp. 1–6, Lisbon, Portugal, November 2013.
- [24] B. Rosario, “Latent semantic indexing: an overview,” Technical report INFOSYS 240 Spring Paper, pp. 1–16, University of Tennessee, Knoxville, TN, USA, 2000.
- [25] J. Wan, D. Wang, S. C. H. Hoi et al., “Deep learning for content-based image retrieval,” in *Proceedings of ACM International Conference 2014*, New York, NY, USA, February 2014.
- [26] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, “Imagenet: a large-scale hierarchical image database,” in *Proceedings 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, USA, June 2009.



Hindawi

Submit your manuscripts at
www.hindawi.com

