IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 19 June 2023, accepted 15 August 2023, date of publication 22 August 2023, date of current version 29 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3307464

RESEARCH ARTICLE

# Link-Driven Study to Enhance Text-Based Image Retrieval: Implicit Links Versus Explicit Links

**KARIM GASMI** [1,2], **HATEM AOUADI** [2], **AND MOUNA TORJMEN** [2]
[1]Department of Computer Science, College of Arts and Sciences at Tabarjal, Jouf University, Jouf, Saudi Arabia
[2]ReDCAD Laboratory, ENIS, University of Sfax, Sfax 3038, Tunisia

Corresponding author: Mouna Torjmen (torjmen.mouna@gmail.com)

**ABSTRACT** Due to the explosive growth of digital images, new efficient and effective methodologies and tools are needed in the image retrieval field. Compared to the content-based image retrieval approach that suffers from the semantic gap, the text-based image retrieval approach has demonstrated its efficiency in retrieving relevant images for a given query. However, this approach suffers from some limitations. For example, query keywords could not match to the textual content of the document or only some images in a document are relevant to the given query. Therefore, a major challenge of the text-based approach is how to improve the image retrieval accuracy without using the image itself, i.e., by using the surrounding information (context) such as the document structure, the links, etc. To achieve this challenge, some works proposed to explore hyperlinks (explicit links) between documents to re-rank images, while more recent works proposed to automatically build implicit links between images and exploit them in the retrieval process. The aim of this paper is thus to compare the exploration of implicit links versus explicit links, either in image ranking or re-ranking. The Image CLEF 2011 collection on Wikipedia shows that not all top-ranked results are interesting to create and analyze linkages between images. In fact, only the aggregate ranking metric makes notice of the fact that linkages improve image retrieval. We also discover that the retrieval strategy–text-based retrieval with no links, implicit link-based re-ranking, or explicit link-based re-ranking–has a significant impact on the efficiency of the query process.

**INDEX TERMS** Implicit links, explicit links, re-ranking, context-based image retrieval.

## I. INTRODUCTION

### A. MOTIVATION

This The immense use of digital camera and social media has led to an exponentially growing of the amount of images uploaded to the Web each day. Consequently, the task of finding relevant images in response to a user's need becomes increasingly difficult. Hence, efficient search algorithms are needed to deal with large amounts of data. The first efforts to improve image retrieval accuracy were concentrated on content-based image retrieval (CBIR) techniques. However, despite the great number of works belonging to this approach, the most challenging task is still closing the semantic gap, defined as "the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation" [1]. Due to the success of web image retrieval methods [2], [3] that are mainly based on keywords, researchers are increasingly focusing on the use of textual information surrounding the images in the retrieval process [4], [5], [6], [7], [8], [9]. Thus, the majority of current image retrieval systems are based on text processing. Unfortunately, the textual information may not describe the real content of the image. Moreover, even if the textual information is relevant to the image, it may not contain the query keywords. Hence, several works [10], [11], [12], [13], [14], [15], [16], [17], [18], [19] have been proposed to improve the effectiveness of text-based image retrieval systems, such as relevance feedback [20], document structure [21], [22], query expansion [12], [23], [24], [25], [26], etc. We are interested in this paper to the use of the links for image re-ranking. The hyperlinks' structure is a good source of evidence in Web information retrieval.

The associate editor coordinating the review of this manuscript and approving it for publication was Anandakumar Haldorai.

In fact, in Web link analysis, the quality of a Web page is generally measured according to the quality of its neighbors, since the underlying assumption is that hyperlinks convey recommendations. However, hyperlinks may not reflect the content similarity between the interconnected web pages due to the existence of spam links (arbitrary created links, navigation links, advertising links, etc.). For this reason, some works [27], [28], [29] propose to create implicit links between documents. However, due to the complexity of computing the semantic similarity between all the documents, several propositions are not evaluated in the information retrieval area. In addition, for image retrieval, most of the works [30], [31], [32], [33], [34] are based on visual features to create implicit links between images. For these reasons, authors in [35] and [36] have proposed, a method for creating implicit links between images through their textual context. This approach proceeds by defining different textual representations (regions) for each image in the collection, and then applying the LDA topic model [37] to compute the similarity between the regions and create the implicit links between images.

### B. CONTRIBUTIONS

In this paper, we aim to compare the use of implicit links versus the use of explicit links (manual hyperlinks) in improving image re-ranking. This comparison is realized using the Wikipedia collection of ImageCLEF 2011 [38] which contains images and hyperlinks between documents. In fact, this collection was adapted to release our link driven study for context-based image retrieval. The main contributions of this work are two folds: 1-compare between building and analyzing explicit (manual) and implicit (automatic) links between images in the retrieval process. 2-evaluate the image re-ranking accuracy when using explicit/implicit links in addition to text-based retrieval.

### C. ARTICLE ORGANIZATION

The remainder of this paper is organized as follows. Section II outlines some related works on analyzing explicit links in text-based image retrieval and on building and analyzing implicit links in image retrieval. We describe in Section III how to create explicit and implicit links between images. The link analysis algorithms used in our study are briefly detailed in Section IV. Experiments and their analysis are conducted and illustrated in Section V, followed by a conclusion and further research directions in Section VI.

## II. RELATED WORK

This section aims to review works using links (either explicit or implicit) in image retrieval.

### A. EXPLICIT LINKS (HYPERLINKS) IN IMAGE RETRIEVAL

Explicit links are hyperlinks between documents that are created by the author of the document. The rich, dynamic structure of these hyperlinks on the Web is not only a help for people browsing the web, but also a powerful help for search engines to understand the relationships between documents and rank these documents more efficiently. Thus, several research works exploiting the hyperlinks for information retrieval were proposed notably within the framework of the Web. Although several works exploiting links have been developed for textual information retrieval, few have been proposed for image retrieval. In most works, classic algorithms such as HITS [39] and PageRank [40] are applied to image search [22]. Dunlop [41] proposed to use links to compute representations for multimedia (non-textual) nodes. Indeed, each multimedia node is assigned a representation composed of the textual contents of all the nodes directly linked to it. The proposed algorithm can be extended to take into account not only the immediate neighbors of a node in the hypermedia network but also neighboring nodes obtained by following two links (a two-step link). The representations will then be used to allow the direct search of these nodes via a textual request. Based on this model, [42] has tried to combine the textual content and the hyperlink structure for multimedia information retrieval on the Web. The representation of an image is composed of sections, and each one has been assigned a weight indicating its importance in the search. Experiments have shown that assigning a higher score to the image caption section (caption) and to the text section of documents linked by two link steps improves the efficiency of the search. Authors in [43] proposed the PicASHOW system based on the application of co-citation-based approaches and methods inspired by PageRank. The assumptions behind the use of co-citation are that: (1) images that are co-contained on the same page are likely to be related to the same subject; (2) images that are contained in pages that are co-cited by other pages are probably related to the same subject. In addition, like PageRank, the PicASHOW system assumes that images that are contained in authoritative pages are good candidates for quality images for the query. The proposed PicASHOW system has improved the performance of web image search. However, PicASHOW supports only keyword queries and cannot handle image content or example image queries. To solve this problem, [44] and [45] proposed WPicASHOW (weighted PicASHOW), allowing a weighted ranking of co-citation analysis that is based on the combination of textual and visual content to regulate the influence of links between pages. The experiments showed a better performance of WPicASHOW over PicASHOW. Authors in [46] and [47] also proposed a system of grouping and retrieving multimedia information on the Web, named iFind. Using a vision-based page segmentation algorithm (VIPS [48]), a web page is partitioned into blocks, and the text and link information of an image can be accurately extracted from the block containing that image. Textual information is used for indexing images. By extracting page-to-block, block-to-image, and block-to-page relationships through the link structure and layout analysis, an image graph is constructed. This method is less sensitive to noise links than previous methods such as PageRank, HITS, and PicASHOW, so the image graph can better reflect the semantic relationships between images. The graph can be used to calculate an importance score for each image, which will then be combined with the relevance score to produce the final rank.

### B. IMPLICIT LINKS IN IMAGE RETRIEVAL

In order to improve the accuracy of image retrieval, several works propose to construct implicit links based on visual content. Thanks to these implicit visual links, a visual graph is constructed and then analyzed to calculate image relevance scores for a given query. To analyze the constructed graph, a random walk method (such as PageRank) has been widely adopted [30], [31], [33], [34], [49]. Authors in [50] proposed a query-specific fusion approach based on graphs, where multiple lists of search results from different visual cues were merged and clustered by link analysis on a merged graph. Authors in [51] have incorporated several visual features in a graph-based learning algorithm for image retrieval. Authors in [32] proposed to construct an off-line visual graph by taking each image as a query and then making links between the query image and the top k returned images. The HITS algorithm is then applied to the set of images returned at querying time. In the same way, authors in [52] followed the same offline step, and at query time, they merged the different graphs obtained using different descriptors. Then, they applied a local ranking algorithm to the resulting graph. For the first time, the creation of implicit links using textual information is taken into account in [53], where the authors have built a multilayer graph where each layer represents a modality (textual, visual, etc.). The constructed graph is undirected, where each node is connected only with its k-nearest neighbors (in terms of Euclidean distance). Then, they applied a random walk algorithm to the multilayer graph by making transitions between the different layers. The proposed solution achieved good image retrieval performance compared to state-of-the-art methods. In [54], the authors utilize hypergraph instead of ordinary graph to model social images. The hypergraph built on the Flickr image dataset contains three types of vertices (i.e., users, tags, and images) and three types of relations (i.e., tagging relations on images, comment relations on images, and similarity relations on images). Based on the hypergraph, they propose HIRT system, which uses Personalized PageRank algorithm to measure vertex similarity, and employs top-k search to support image retrieval and tagging. Experimental analysis on a large Flickr dataset shows the effectiveness and efficiency of this system compared with existing system and techniques. In [55] a fully-connected graph is constructed using the database of images and the process of ranking a set of items based on a single set of scores is defined as a problem of finding a suitable path in this graph. In this graph, an edge weight is equivalent to the absolute difference between the similarity scores (with respect to the query) of the two images corresponding to the two nodes that it connects. Authors in [56] demonstrate the necessity of graph fusion in multifeature image retrieval task. After a K-nearest neighbors search of a query, the Jaccard coefficient is adopted to initialize the weight of the constructed neighborhood graph, then a multigraph fusion ranking method is proposed. Reference [57] proposed a graph-based selective rank fusion for image retrieval. In the constructed graph, features are represented by nodes, while the edges represent the complementarity relationships. The graph is iteratively updated according to different thresholds of effectiveness estimation measures, which aims at selecting the most effective features. The selected features are defined by connected components of the graph, which are combined by a late fusion approach for final ranking. High-effective retrieval results have been obtained, yielding relative gains up to +54.73% in relation to the best-isolated feature. Some other works have been done as part of the MediaEval [58], [59], [60], [61], [62], [63] and TRECVID [61], [64] evaluation campaigns for video hyperlinking. Authors in [61] concluded that textual features work better than visual ones in this task, whereas visual features by themselves cannot predict reliable hyperlinks. Nevertheless, they suggest that the use of visual features to rerank the results of text-based retrieval can improve performance. A more recent work [35], [36] proposed to use the LDA model to construct textual representations for each image in the collection, and then compute similarities between these images' representations to decide whether or not to create a link between two images. The experiments showed the interest in this approach. In conclusion, several works in the literature have explored and studied the efficiency of link analysis in image retrieval. But to our best knowledge, no studies have compared the exploration of implicit versus explicit links between images.

## III. EXPLICIT/IMPLICIT LINK BUILDING BETWEEN IMAGES

In this section, we describe the approaches for creating respectively explicit and implicit links between images of the collection in our study.

### A. BUILDING EXPLICIT LINKS BETWEEN IMAGES

To our knowledge, there is no available collection aiming to evaluate the use of hyperlinks between images in the context of the web (context-based image retrieval). In the ImageCLEF Wikipedia 2011 collection, there are hyperlinks between documents, i.e., a text part is pointing to an entire document, but not hyperlinks between images. We propose to propagate the hyperlinks between documents and images as follows: -bidirectional links are created between images belonging to the same document (continuous arrows in Figure 1). -a link from a document d1 to another document d2 means the creation of links from all images of d1 to all images of d2 (discontinued arrows in Figure 1). An example of an explicit link-based graph between documents and an explicit link based graph between their associated images is given in Figure 1.

In the part (a), there are three nodes (documents) and three edges (hyperlinks). After applying the hyperlink creation

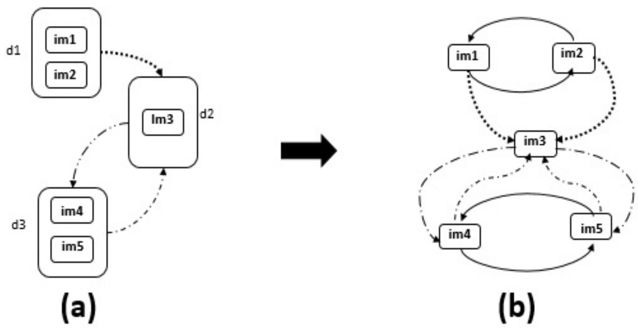


**FIGURE 1.** Explicit link graph between documents vs Explicit link graph between associated images.

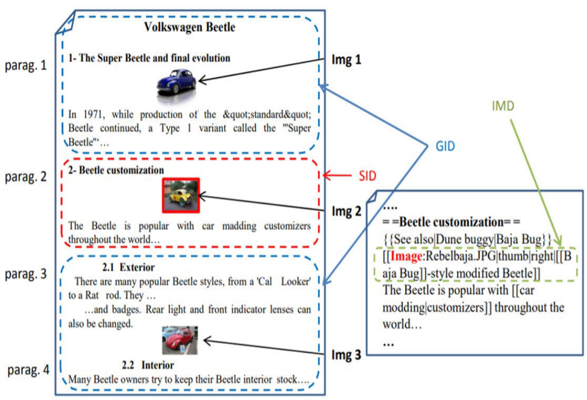


**FIGURE 2.** MID, SID and GID representations of an image [36].

process between images, we obtain (part (b)), 5 nodes (images) and 10 edges (hyperlinks).

### B. BUILDING IMPLICIT LINKS BETWEEN IMAGES

This section describes the work proposed in [36], which aiming to automatically create implicit links between images through computing the similarity of the textual information representing them. The first step is to segment the document content into three representations for each contained image: "Specific Image Description" (SID), "Generic Image Description" (GID), and "Image Metadata" (IMD). The IMD representation is the image caption; the SID representation is the textual paragraph containing the image; and the GID representation is the document's textual content except the textual paragraph containing the image. Figure 2 illustrates the three representations of an image.

The second step is to create implicit links between images via the LDA (Latent Dirichlet Allocation [37]) topic model. In fact, documents are presented by topic distributions, which are used to calculate semantic similarities between images. In [65], authors have shown the efficiency of this model in context-based image retrieval by applying it to retrieve relevant images without taking into account the links between documents. This process is applied separately for each image representation (IMD, SID, and GID). Figure 3 presents an overview of the implicit link creation process for two images (Image i and Image j)

As illustrated in the figure, after performing the LDA process, each image is represented by a topic distribution. In order to create the links between images, a similarity score is computed between both distributions, and an adapted version of the classical cosine measure is used. The new similarity measure includes the number of common topics between the two image representations as follows:

$$* \text{ sim}\left(\overrightarrow{Rep_{img_i}}, \overrightarrow{Rep_{img_j}}\right) = \cos_{sim}\left(\overrightarrow{Rep_{img}}, \overrightarrow{Rep_{img_j}}\right) * \left|\text{commonTopics}\left(\overrightarrow{Rep_{img}}, \overrightarrow{Rep_{img_j}}\right)\right| \quad (1)$$

$\overrightarrow{Rep_{img_i}}$ (respectively $\overrightarrow{Rep_{img_j}}$) is the topic distribution of the representation $Rep_{img_i}$ resp. $Rep_{img_j}$ Finally, a threshold is applied to remove links having low scores and reduce the number of created links. This similarity threshold is set to 0.1 according to [36]. Figure 4 illustrates the application of a threshold equals to 0.1 on a SID graph.

To determine the link's direction, the following formula, based on the percentage of shared information between the two representations of the image, is used:

$$\text{Score}_{\text{Direction}} \; (Rep_{img_i} \rightarrow \text{Rep} \quad (2)$$

With $Rep_{img_i} \rightarrow Rep_{img_j}$ means that the links is from $Rep_{img_i}$ to $Rep_{img_j}$. Algorithm 1 presents the process of building implicit links.

---

**Algorithm 1** Building Implicit Links Between Images

---

**Require:** Document collection
**Ensure:** Graph of linked images
  Document segmentation into SID, GID and IMD
  **for** each image representation $Rep_{img}$ (IMD, SID, and GID) **do**
    Apply LDA to extract topic distributions ($\overrightarrow{Rep_{img}}$)
  **end for**
  **for** each $Rep_{img}$ **do**
    **for** each pair $\overrightarrow{Rep_{img_i}}$ and $\overrightarrow{Rep_{img_j}}$ **do**
      Compute similarity using (1)
      **if** $sim(\overrightarrow{Rep_{img_i}}, \overrightarrow{Rep_{img_j}}) > 0.1$ **then**
        Apply (2) to add a weighted directed edge between two images $img_i$ and $img_j$
      **end if**
    **end for**
  **end for**

---

The last step is analyzing the created links to determine the image scores. A link-based score for each image representation is thus calculated. Therefore, each image is assigned three link-based scores according to the image representation: the SID link score, the GID link score, and the IMD link score, as illustrated in part (a) of Figure 5.

Finally, to obtain a single link-based score, the three scores are combined as follows:

$$\text{Link}_{\text{Final}} = \alpha * \text{Lin}_{LMD} + \beta * \text{Lin}_{SID} + \gamma * \text{Link}_{GID} \quad (3)$$

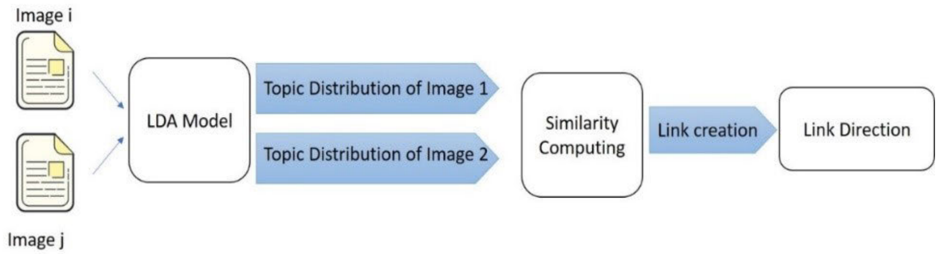


**FIGURE 3.** Overview of the LDA based link creation approach between two images.

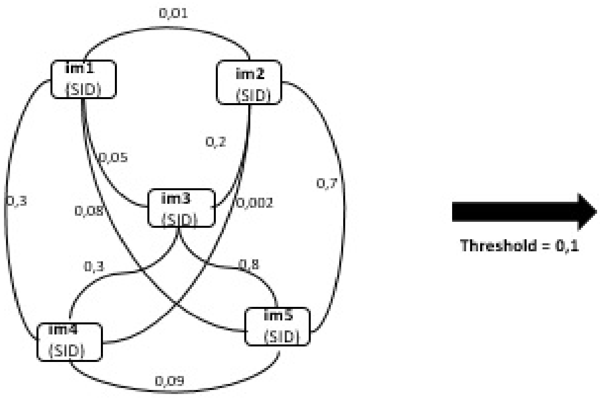


**FIGURE 4.** Application of a threshold to reduce the number of implicit links.

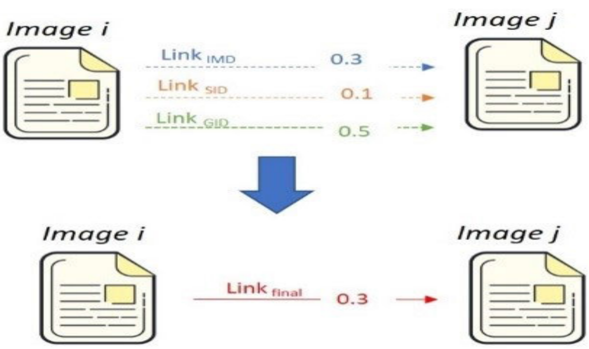


**FIGURE 5.** Example of linking images according to the SID, GID and IMD image representations.

where $\alpha$, $\beta$ and $y$ are parameters used to adjust the participation of each score in the final link score. Their sum quals to $\infty$, In this paper, we set $\alpha = \beta = \gamma = 1/3$ (part (b) of Figure 5). The LDA parameters were fixed as follows: $\alpha = 50/\mathrm{K}$ where $\mathrm{K} = 1500$ and $\beta = 0.01$ according to [36].

## IV. LINK ANALYSES ALGORITHMS

For link analysis, we propose to use three types of measures: a geometric measure (Degree Centrality [66]), a path-based measure (Betweenness Centrality [67]) and a spectral measure (HITS [39]). In the following, we briefly detail the three measures.

### A. DEGREE CENTRALITY

Degree Centrality is defined as the number of links incident upon a node. This measure is considered the predecessor of all measures used in link analysis, as it used a simple heuristic to rank the nodes according to their popularity. In the case of a directed network (links have direction, such as on the web), two separate measures of degree centrality are defined: indegree and outdegree. Accordingly, indegree is a count of the number of links directed to the node, and outdegree is the number of links that the node directs to others. Indegree is often interpreted as a form of popularity, while outdegree is interpreted as a form of gregariousness. In our case, we are interested in the degree of centrality (DCIN), which is defined as follows:

$$DC_i^{IN} = \sum_{j=1}^{n} e_{ij} \tag{4}$$

with $i$ and $j$ are two nodes; $e_{ij}$ is a link directed from $j$ to $i$; $n$ is the number of links pointing to the node $i$.

### B. BETWEENNESS CENTRALITY

Betweenness Centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. The Betweenness Centrality (BC) measure is defined as follows:

$$BC_i = \sum\nolimits_{j \to k} \frac{g_{jk}(i)}{g_{jk}} \tag{5}$$

where $g_j k$ is the number of shortest paths (according to the number of nodes in the path) from node $j$ to node $k$, and $g_{jk}(i)$ is the number of those paths passing through the node $i$

### C. HYPERLINK INDUCED TOPIC SEARCH

The basic idea behind the HITS (hyperlink-induced topic search) algorithm is to build a query-specific graph, called a ''neighborhood graph,'' which contains only pages (nodes) related to the query (HITS is a query-dependent algorithm). These pages are divided into (1) pages that provide trustworthy information on a given query, called authorities, and (2) pages that contain links to authorities, called hubs. Authorities and hubs exhibit a mutually reinforcing relationship: a better hub points to many good authorities, and a better authority is pointed to by many good hubs. The authority score A of a node i is the sum of all hub scores H of nodes that point to i. The hub score H of a node i is the sum of all authority scores A of nodes that are pointed by i, as shown in the following equation:

$$A_i = \sum_{j \in G_i} H_j \quad \text{and} \quad H_i = \sum_{j \in G_i} A_j \tag{6}$$

where $G_i$ in the set of nodes in the neighbourhoed graph of the node $l$ ;$j$ is a node belonging to $G_i$ The final authority and hub scores of a node are obtained through an iterative application of the equation until convergence. Nedes are then canked by the authelity scere.

## V. EXPERIMENTS AND DISCUSSION

In this section, we begin by describing the experimental setup, then we present our results and interpretations of our comparison study between exploiting implicit and explicit links in image retrieval. Finally, we discuss our findings and present some future issues.

### A. EXPERIMENTAL SETUP

In this section, we first present the dataset used for the comparison between the use of implicit and explicit links in context-based image retrieval. Then we detail the metrics used to evaluate our experiments, and finally, we describe how the textual run, which is the baseline, is done.

#### 1) DATASET

To our knowledge, there is no collection composed of images linked between them by implicit or explicit links. Thus, we propose to use the Wikipedia collection provided by the ImageCLEF (CLEF Cross Language Image Retrieval Track) evaluation campaign in 2011 for the Wikipedia retrieval task [38]. The dataset was composed of 125 827 documents extracted from Wikipedia in three languages, containing 237 434 images. A set of 50 queries was provided to perform the retrieval accuracy evaluation. This dataset contains hyperlinks (explicit links) between documents but not between images. We have already described in Section III-A how to build explicit links between images by using explicit links between documents. For the technical limitations of the memory requirement to build a multilingual topic model when creating implicit links between images, we have used in our experiments only documents written in English where the number is 42 774. However, our approach can be applied to any language and any type of document. For queries, we have used the keywords of the English title without the small set of example images provided with each textual query. In order to properly evaluate our proposition, we have constructed for each query a new base of assessments composed only of images belonging to the English documents. Figure 6 presents an extract of the original base of assessments (document-based assessments) and an extract of the new base of assessments (image-based assessments).

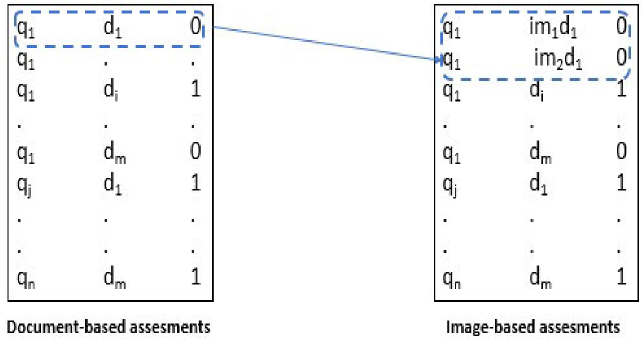


**FIGURE 6. The construction process of image-based assessments.**

-$q_i$ is the query number with $n$ is the number of queries.

-$d_i$ is the document identity with $m$ is the number of documents.

-$im_k d_i$ is the image position $k$ in the document $d_i$

-The last column in the assessments presents the relevance value: 0 for not relevant and 1 for relevant.

According to Figure 1, $d_1$ is not relevant for query $q_1$ in original assessments and since $d_1$ contains two images $im_1 d_1$ and $im_2 d_1$, then both images are not relevant in the constructed assessments.

#### 2) MEASURES

For the evaluation, we used early precision (P@X) and mean average precision (MAP) measures. In fact, P@X presents early precisions (P@5, P@10 ...) which are relevant in the Web search context, since users in general examine relatively the top ranked results.

Whereas, the MAP measure is used to evaluate the global system's effectiveness [68]. The precisions are calculated for each relevant retrieved document. The average precision (AP) reflects the query performance. The MAP is the average of the average precisions for all queries and is obtained with the following formula:

$$MAP = \frac{\sum_{q \epsilon Q} AP_q}{|Q|} \tag{7}$$

where APq is the average precision of a query q, Q is the query set, and |Q| is the number of queries. To validate our results statistically, we used the Wilcoxon signed-rank test [69], which is the non-parametric equivalent of the paired-samples t-test. This test consists in calculating the value of significance $p \in [0, 1]$, which estimates the probability that the difference between two methods is due to chance. We can conclude that the two methods are statistically different (and not by chance) when $p < \alpha$, where $< 0.05$ is commonly used [70]; the nearer $p$ approaches zero, the more likely it is that the two methods are different. In our experiments, we consider that the difference between two methods is significant if $p < 0.05$.

#### 3) COMPUTING TEXTUAL SCORES (BASELINE)

Due to computing complexity and time-space costs in link analysing, we decided to run a textual search and then apply the link analysis for only the first 1000 returned documents. We should also mention that building implicit links is a very intensive process in terms of computation and memory. To compute the textual based scores of images, the basic vectorial model [71] was used.

### B. RESULTS AND ANALYSIS

In this section, we first compare the use of implicit and explicit links according to link scores and without using textual scores. Then, we compare the re-ranking process of images by combining textual scores and link scores (once with implicit links and once with explicit links). Finally, we interpret and explain the different results.

#### 1) IMPLICIT VS. EXPLICIT LINKS IN IMAGE RE-RANKING WITHOUT TEXTUAL SCORES

This section aims to compare the effectiveness of using implicit links against explicit links in image re-ranking using only link-based scores of images. We talk here about re-ranking because we first run a textual model to get the best 1000 documents relevant to the query, and then we use these 1000 documents to analyze links between them and re-rank results. Table 1 depicts the overall results, where scores obtained only by analyzing links are used to rerank images. Link analysis algorithms are indicated in the first column. The link type used in ranking, i.e., explicit (Exp) or implicit (Imp), is indicated in the first column. The run name is obtained by combining the first letters of the link analysis method and the link type, e.g., DCIL is Degree Centrality Implicit Links. Difference rates (% Diff/Imp) are computed between implicit and explicit links. Starred differences mean that there is a significant improvement of using explicit links compared to implicit links according to Wilcoxon test (p<0.05).. According to the results, explicit links-based re-ranking greatly outperforms implicit links-based re-ranking according to the three link analysis algorithms and to the three measures (difference rates are between 20% and 81% and are statistically significant: p value<0.05). These results are not surprising since the links are created manually and the collection used (Image Clef Wikipedia Retrieval Task 2011) is a clean dataset that contains only informational links between documents, i.e., navigational and marketing links are removed. Consequently, the hyperlinks in this collection reflect a high degree of semantic similarity between documents. However, in real Web collection, documents contain usually noisy links that could be used in the re-ranking process and affect negatively the results. More experiments with real web documents without cleaning processes should be released to evaluate the efficiency of using hyperlinks in image retrieval. Comparing the use of the different link analysis algorithms, Table 1 shows that degree centrality leads to the best accuracy for both implicit and explicit links according to all metrics. These results can be argued by the assumptions behind the processing of each algorithm. Thus, in Betweenness Centrality, the nodes are considered important if they connect other nodes, regardless of the number of connections they have. This measure can be seen as the potential of a node to be a broker between two or more groups of nodes. Although this measure seems to be very interesting in social networks since it has a large influence on the control of the flow in a graph (if a broker node is removed, the flow will be stopped), It is not the case in the information retrieval and re-ranking processes. In fact, the aim of creating a link between two documents in information retrieval is not to transfer information from one node (document) to another, but to reference a node as relevant information to the source one. Regarding the Degree Centrality measure, it assumes that nodes with more connections (the number of edges that they have) are more influential and important in a network. This measure seems interesting for information retrieval since a document with many links will be considered a good reference (an important node in the graph). On the other side, the HITS algorithm assumes that nodes can be important for the network, but they are not central. In order to find such nodes, the HITS algorithm introduces two types of central nodes: hubs and authorities. Authorities are the most cited nodes by hubs, and hubs are the nodes citing the highest authorities. Taking into account that HITS is a query-dependent algorithm, the main idea is to select the K top-ranked documents according to the query and then extend this initial set of roots to other documents. This basic idea is not respected in our work, as we have used only the top 1000 documents ranked by a textual model without the extension step.

#### 2) IMPLICIT VS. EXPLICIT LINKS IN IMAGE RE-RANKING WITH TEXTUAL SCORES

In this section, we have done the same experiments as in the previous section, but by combining link-based scores and text-based scores to re-rank images. We have also displayed the results of the textual run in order to study the impact of using links as an evidence source compared to using only text information. Best values of each metric are bold. Difference rates between implicit/explicit links runs and textual run are computed. Significant differences according to the Wilcoxon test are starred (p<0.05). According to all measures, we note that explicit links did not improve results compared to the use of text only. No significant improvement is noted (difference rates with the triangle-right symbol). Doing the same comparison using implicit links (Difference rates with ♣ symbol), we note that there is slight difference that is not significant at P@5, but a significant improvement is observed at P@10 and MAP measures. Moreover, the Degree Centrality algorithm provided the best performance in image re-ranking, as shown

**TABLE 1.** Results of image re-ranking without textual scores.

| Link Type | Algorithm | Run Name | P@5 | %Diff/Imp | P@10 | %Diff/Imp | MAP | %Diff/Imp |
|---|---|---|---|---|---|---|---|---|
| 3*Imp | Degree Centrality | DCIL | 0.148 | — | 0.136 | — | 0.094 | — |
| | HITS | HIL | 0.108 | — | 0.104 | — | 0.081 | — |
| | Betweenness Centrality | BCIL | 0.104 | — | 0.084 | — | 0.062 | — |
| 3*Exp | Degree Centrality | DCEL | **0.192** | **+30%*** | **0.164** | **+21%*** | **0.113** | **+20%*** |
| | HITS | HEL | 0.172 | **+59%*** | 0.158 | **+52%*** | 0.113 | **+40%*** |
| | Betweenness Centrality | BCEL | 0.188 | **+24%*** | 0.152 | **+81%*** | 0.096 | **+55%*** |

**TABLE 2.** Results of image re-ranking with textual scores.

| 2*Link Type | 2*Algorithm | 2*Run | 2*P@5 | %Diff | | 2*P@10 | %Diff | | 2*MAP | %Diff | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | /txt | /Exp | | /txt | /Exp | | /txt | /Exp |
| Textual | None | Txt | 0.384 | - | - | 0.316 | - | - | 0.239 | - | - |
| 3*Exp+text | Degree Centrality | DCELT | 0.392 | +2% (▷) | - | 0.334 | +6% (▷) | - | 0.243 | +2% (▷) | - |
| | HITS | HELT | 0.376 | -2% (▷) | - | 0 .320 | +1% (▷) | - | 0.242 | +1% (▷) | - |
| | Betweeness Centrality | BCELT | 0.388 | +1% (▷) | - | 0.320 | +1% (▷) | - | 0.240 | 0% (▷) | - |
| 3*Imp+text | Degree Centrality | DCILT | **0.400** | **+4%** (♣) | +2% (♠) | **0.352** | **+11%** (♣) | **+5%** (♠) | **0.270** | **+13%** (♣) | **+11%** (♠) |
| | HITS | HILT | **0.400** | +4% (♣) | +6% (♠) | 0.338 | +7% (♣) | +6% (♠) | 0.260 | +9% (♣) | +8% (♠) |
| | Betweeness Centrality | BCILT | 0.396 | +3% (♣) | +2% (♠) | 0.338 | +7% (♣) | +6% (♠) | 0.267 | +11% (♣) | +11% (♠) |

in the Table 1 results. However, in contrast to the results of image ranking based only on link scores (Table 1), we found that the use of implicit links was better than the use of explicit hyperlinks (difference rates with the ♠ symbol) for image re-ranking since it provided a significant improvement according to the three link analysis algorithms (e.g., +4% for P@5, +5% for P@10, and +11% for MAP when using Degree Centrality). These results attract attention, but they could be explained by the relation between the relevance density and the link density. The relevance density is defined by the proportion of the returned documents that contain relevant documents, while the link density is defined by the proportion of the returned interlinked documents.

### 3) INTERPRETATION OF RESULTS

In our case, regarding the explicit link density (number of hyperlinks) on the list of returned images, it is obvious that this density is very high in the top-ranked images since the textual representations of these top-ranked images share the most common words with each other thanks to the common terms in the query. More precisely, the textual representations of the top ranked images share with the query more common terms than the down ranked images. This explanation is illustrated in part (a) of Figure 7. Moreover, the hyperlink density (explicit links) is higher in the top-ranked images than the lower-ranked ones since there are more common terms between the textual representations of images. Part b of Figure 7 illustrates this interpretation. However, for implicit links, the density is calibrated and scattered along the entire list since the link creation is based on the topic modeling. In fact, the top-ranked images could be linked with the lower-ranked images if they share the same topic, even if there are few common terms between their textual representations. This explanation is illustrated in part c of Figure 7.

Since the relevance density is generally very high in the top ranked images for a given query and the hyperlink density is also very high in the top ranked images, the re-ranking accuracy was slightly affected when combining textual scores with link-based ones, which argues that the combination did not have a positive impact on the retrieval accuracy. However, implicit links analysis was interesting for image re-ranking thanks to the LDA model: several returned images have changed their rank positively thanks to the implicit links.

Finally, releasing some statistics on both link types in the collection, we have found that only 7% of implicit links (automatically created) are also explicit links, and 10% of explicit links (manual hyperlinks) are also implicit links. These statistics show that there is a gap between both link-types, and we plan to study combining both link-types in image re-ranking to enhance accuracy.

## C. DISCUSSION AND ISSUES

According to the previous experiments, the main findings of our study could be summarized as follows: 1- Creating implicit links between images seems like a good solution for improving text-based image retrieval accuracy since combining the textual scores with the implicit links scores allows for a significant improvement in accuracy.

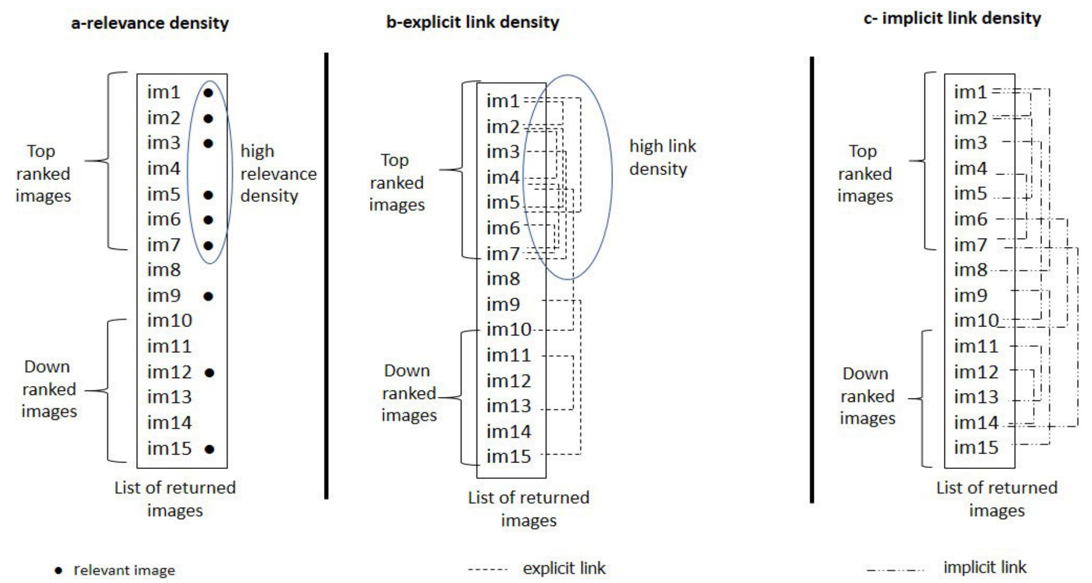


**FIGURE 7.** Relevance density vs link density throughout returned documents for a given query.

2- The benefit of the link analysis on the image re-ranking depends on two features: the nature of the links between documents (purely semantic or containing navigational and advertising links) and the link density. In fact, explicit links outperformed implicit links when ranking images only by link scores (Table 1), whereas the contrary was obtained when combining link scores with textual ones (Table 2).

3- The combination of both link types could enhance furthermore the image re-ranking accuracy since a gap is observed when overlapping both types of links: explicit links created by the authors are more semantic and accurate whereas implicit links created automatically covers more documents.

To better understand the impact of using links on image re-ranking, we have released query-by-query analysis and computed the number of queries having the best results according to each link strategy: using implicit links, using explicit links, and without using links, according to the following measures: P@5, P@10 and MAP. Table 3 presents the statistics of the used queries in our study.

According to Table 3, we notice that 72% of queries have the best results without the use of links (36 queries out of a total of 50) for P@5. By the same way, 56% of queries obtained the best results when there was no use of links (28 queries among a total of 50) for P@10. These results show that links are not very beneficial for enhancing image re-ranking in top results and that using only textual information is sufficient. These results are not very surprising, as the top-ranked documents generally contain all the words of the given query, and consequently, they have a high probability of being relevant. However, when analysing the overall performance (according to MAP measure), we note that 82 % of queries have obtained the best results when the links are used (14+27 queries among a total of 50). In addition, we can observe that the use of implicit links is better than the use of explicit links (9 vs. 5 for P@5, 12 vs. 10 for P@10, and 27 vs. 14 for MAP, respectively), which confirms our assumption that building and analyzing implicit links for image retrieval is more beneficial than creating and using explicit links (hyperlinks). In the last part of our study, we constructed an artificial run (named in our experiments "Best Comb") by keeping the best result over the three strategies (implicit links vs. explicit links vs. no links) for each query according to the MAP value. So, in the Best Comb run, 14 queries are processed by the explicit links, 27 queries are processed by the implicit links, and 9 queries are processed only by textual information. Table 4 presents the results of the textual run, the best implicit links-based run, the best explicit links-based run, and the best comb run.

The main conclusion of Table 4 is that the best results are obtained by the "best comb" run, which is a combination of the three link strategies. which means that for some queries in text-based image retrieval, it is better to create and analyze implicit links; for other queries, it is better to create and analyze explicit links; and for a third set of queries, it is better to not use links outright. Based on our study and our findings, a new challenge appears: how to classify textual queries to decide the best strategy to search for relevant images: text-based image retrieval without links, implicit links-based image re-ranking, or explicit links-based image re-ranking? By examining the set of queries used in our study, we think that some query features could be interesting in the classification process, such as the query size, the existence or not of a color term, an entity name, or a phrase. Moreover, semantic relatedness between query terms could also be a leading feature for the query classification.

At the end, we confront our findings with the results of the work presented in [65] that did not use links (neither implicit

**TABLE 3. Query set statistics according to best results.**

| | Explicit links | Implicit links | Without links (txt) | Total Queries |
|---|---|---|---|---|
| P@5 | 5 | 9 | 36 | 50 |
| P@10 | 10 | 12 | 28 | 50 |
| MAP | 14 | 27 | 9 | 50 |

**TABLE 4. Results of the different strategies in our link driven study for image retrieval.**

| Strategy | Run | P@5 | %Diff /txt | P@10 | %Diff /txt | MAP | %Diff /txt |
|---|---|---|---|---|---|---|---|
| Txt/No Links | Txt | 0.384 | - | 0.316 | - | 0.239 | - |
| Explicit Links | DCELT | 0.392 | +2% | 0.334 | +6% | 0.243 | +2% |
| Implicit Links | DCILT | 0.400 | +4% | 0.352 | +11% * | 0.270 | +13% * |
| Combination of the three previous strategies | Best Comb | **0.448** | **+17%*** | **0.378** | **+20% *** | **0.311** | **+30% *** |

**TABLE 5. Comparison between our work and literature works.**

| | run | P@5 | P@10 | MAP |
|---|---|---|---|---|
| Text + Implicit Links | DCILT | 0.400 | 0.352 | 0.270 |
| Similar work without links | PAR_LDA | 0.400 | 0.336 | 0.281 |
| Combination of the three strategies | Best Comb | **0.448** | **0.378** | **0.311** |

**TABLE 6. Results of approaches using only simple textual information in the Wikipedia retrieval task 2011.**

| Rank | Participant | MAP |
|---|---|---|
| 1 | Best Comb | 0.311 |
| 2 | UNED | 0.304 |
| 3 | DEMIR | 0.243 |
| 4 | ReDCAD | 0.230 |
| 5 | SZTAKI | 0.216 |
| 6 | DBISForMaT | 0.209 |
| 7 | SINAI | 0.206 |

nor explicit). In this work, the authors have proposed a document segmentation method to represent images by a specific textual part of the document, and then we have applied the LDA model to retrieve relevant images for each given query. The results presented in Table 5 show that the work [65] based on a specific method of using textual content of documents to represent images has the same accuracy (0.400 at P@5) and sometimes outperforms (0.281 vs. 0.270 at MAP) the run exploiting implicit links between images. These results confirm the assumption that the links were not always beneficial.

However, we observe clearly that the combination of the three strategies by selecting the best strategy according to the query features improves significatively the accuracy over all the approaches (+12% at P@5; +13% at P@10, +11% at MAP compared to the work [65] with p-value<0.05).

We also compare our work with the official runs in the Wikipedia retrieval task 2011. To do this fairly, we take into account only official runs using textual information without relevance feedback or query expansion. Table 6 shows official results ranked by MAP.

The results show that our work outperforms all the proposed methods in Wikipedia retrieval task 2011.

## VI. CONCLUSION

In this paper, we have conducted a link-driven study for enhancing text-based image retrieval. In fact, we have compared the use of textual information as the only source of evidence to retrieve relevant images for a given query and the exploration of explicit (respectively implicit) links between images to improve image retrieval accuracy. Experimental results demonstrated that the use of implicit links compared to the use of explicit links has 5% more accuracy at top-ranked results and 11% more accuracy at overall accuracy. These improvements are statistically significant. Query-by-query analysis showed that in the top-ranked results, the use of links is not always beneficial and the text around images is a sufficient source of evidence to determine the image relevance for a given query (72% and 56% of queries for P@5 and P@10, respectively). However, according to the overall performance of the image retrieval process, the use of links is very interesting and allows to enhance textual results very well (82% of queries for MAP measure). Another interesting finding of our study is that, overall, the best strategy to retrieve images (no links, implicit links, explicit links) depends on the query, and a good choice of strategy allows for the best accuracy. Based on this finding, a new challenge appeared: define a query classifier to decide which strategy is the best to retrieve relevant images by analyzing the query.

In future work, we plan to study and classify queries according to some features in order to choose the best strategy. In addition, we want to further investigate the combination of explicit and implicit links in image retrieval. On the one hand, explicit hyperlinks are more semantic than implicit links if they are informational since they are created manually, and on the other hand, it is not possible to link manually all similar documents in the collection, so it is interesting to build automatically implicit links between images to take into account all possible similarities. The motivation behind this perspective is that only an overlapping of 10% was found between both implicit and explicit links. Finally, another issue to investigate in our work is multi-modal image retrieval. The combination of content and context-based image retrieval has shown its effectiveness in several works. Therefore, we plan to explore visual features in the construction of implicit links. It is possible, for example, to build visual-based links and explore them in the image retrieval process.

## REFERENCES

[1] 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 Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.

[2] M.-D. Nguyen, B. T. Nguyen, and C. Gurrin, "HADA: A graph-based amalgamation framework in image-text retrieval," in *Proc. Eur. Conf. Inf. Retr.*, Mar. 2023, pp. 717–731. [Online]. Available: https://api.semanticscholar.org/CorpusID:255749506

[3] J. Tekli, "An overview of cluster-based image search result organization: Background, techniques, and ongoing challenges," *Knowl. Inf. Syst.*, vol. 64, no. 3, pp. 589–642, Mar. 2022.

[4] A. Brandsen, S. Verberne, K. Lambers, and M. Wansleeben, "Can BERT dig it? Named entity recognition for information retrieval in the archaeology domain," *J. Comput. Cultural Heritage*, vol. 15, no. 3, pp. 1–18, Sep. 2022.

[5] Y. Fan, X. Xie, Y. Cai, J. Chen, X. Ma, X. Li, R. Zhang, and J. Guo, "Pre-training methods in information retrieval," *Found. Trends Inf. Retr.*, vol. 16, no. 3, pp. 178–317, 2022.

[6] Y. Kim, S. Bang, J. Sohn, and H. Kim, "Question answering method for infrastructure damage information retrieval from textual data using bidirectional encoder representations from transformers," *Autom. Construct.*, vol. 134, Feb. 2022, Art. no. 104061.

[7] R. Kumar and S. C. Sharma, "Hybrid optimization and ontology-based semantic model for efficient text-based information retrieval," *J. Supercomput.*, vol. 79, no. 2, pp. 2251–2280, Feb. 2023.

[8] K. Gasmi, M. T. Khemakhem, and M. B. Jemaa, "A conceptual model for word sense disambiguation in medical image retrieval," in *Proc. 9th Asia Inf. Retr. Societies Conf., Inf. Retr. Technol. (AIRS)*. Singapore: Springer, Dec. 2013, pp. 296–307.

[9] H. Ayadi, M. Torjmen-Khemakhem, M. Daoud, J. X. Huang, and M. Ben Jemaa, "MF-Re-Rank: A modality feature-based re-ranking model for medical image retrieval," *J. Assoc. Inf. Sci. Technol.*, vol. 69, no. 9, pp. 1095–1108, Sep. 2018.

[10] A. Oussama, B. Khaldi, and M. L. Kherfi, "A fast weighted multi-view Bayesian learning scheme with deep learning for text-based image retrieval from unlabeled galleries," *Multimedia Tools Appl.*, vol. 82, pp. 10795–10812, Sep. 2022.

[11] K. Pereira, A. Parikh, P. Kumar, and V. Hole, "Multilingual text-based image search using multimodal embeddings," in *Proc. IEEE 6th Conf. Inf. Commun. Technol. (CICT)*, Nov. 2022, pp. 1–5.

[12] Z. Gong and Q. Liu, "Improving keyword based web image search with visual feature distribution and term expansion," *Knowl. Inf. Syst.*, vol. 21, no. 1, pp. 113–132, Oct. 2009.

[13] M. Torjmen, K. Pinel-Sauvagnat, and M. Boughanem, "Using textual and structural context for searching multimedia elements," *Int. J. Bus. Intell. Data Mining*, vol. 5, no. 4, pp. 323–352, Oct. 2010.

[14] S. Shehata, F. Karray, and M. S. Kamel, "An efficient concept-based retrieval model for enhancing text retrieval quality," *Knowl. Inf. Syst.*, vol. 35, no. 2, pp. 411–434, May 2013.

[15] T. Mei, Y. Rui, S. Li, and Q. Tian, "Multimedia search reranking: A literature survey," *ACM Comput. Surveys*, vol. 46, no. 3, pp. 1–38, Jan. 2014.

[16] R. T. Calumby, M. A. Gonçalves, and R. D. S. Torres, "On interactive learning-to-rank for IR: Overview, recent advances, challenges, and directions," *Neurocomputing*, vol. 208, pp. 3–24, Oct. 2016.

[17] A. Iftene and B. Alexandra-Mihaela, "Using semantic resources in image retrieval," *Proc. Comput. Sci.*, vol. 96, pp. 436–445, Dec. 2016.

[18] P. Jing, Y. Su, C. Xu, and L. Zhang, "HyperSSR: A hypergraph based semi-supervised ranking method for visual search reranking," *Neurocomputing*, vol. 274, pp. 50–57, Jan. 2018.

[19] C. Cui, J. Shen, Z. Chen, S. Wang, and J. Ma, "Learning to rank images for complex queries in concept-based search," *Neurocomputing*, vol. 274, pp. 19–28, Jan. 2018.

[20] X. Wang, C. MacDonald, N. Tonellotto, and I. Ounis, "ColBERT-PRF: Semantic pseudo-relevance feedback for dense passage and document retrieval," *ACM Trans. Web*, vol. 17, no. 1, pp. 1–39, Feb. 2023.

[21] M. Torjmen-Khemakhem, K. Pinel-Sauvagnat, and M. Boughanem, "Investigating the document structure as a source of evidence for multimedia fragment retrieval," *Inf. Process. Manage.*, vol. 49, no. 6, pp. 1281–1300, Nov. 2013.

[22] H. Aouadi, M. Torjmen-Khemakhem, and M. B. Jemaa, "Combination of document structure and links for multimedia object retrieval," *J. Inf. Sci.*, vol. 38, no. 5, pp. 442–458, Oct. 2012.

[23] R. Srivel, K. Kalaiselvi, S. Shanthi, and U. Perumal, "An automation query expansion strategy for information retrieval by using fuzzy based grasshopper optimization algorithm on medical datasets," *Concurrency Comput., Pract. Exper.*, vol. 35, no. 3, Feb. 2023, Art. no. e7418.

[24] M. Torjmen-Khemakhem and K. Gasmi, "Document/query expansion based on selecting significant concepts for context based retrieval of medical images," *J. Biomed. Informat.*, vol. 95, Jul. 2019, Art. no. 103210.

[25] A. P. Natsev, A. Haubold, J. Tešić, L. Xie, and R. Yan, "Semantic concept-based query expansion and re-ranking for multimedia retrieval," in *Proc. 15th ACM Int. Conf. Multimedia*, Sep. 2007, pp. 991–1000.

[26] S. Jeyappratha, S. C. Punitha, and M. Phil, "An efficient image re ranking based on keyword expansion using query log," *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 4, pp. 5935–5938, 2014.

[27] G.-R. Xue, H.-J. Zeng, Z. Chen, W.-Y. Ma, H.-J. Zhang, and C.-J. Lu, "Implicit link analysis for small web search," in *Proc. 26th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jul. 2003, pp. 56–63.

[28] O. Kurland and L. Lee, "Respect my authority!: HITS without hyperlinks, utilizing cluster-based language models," in *Proc. 29th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Aug. 2006, pp. 83–90.

[29] A. Belmouhcine and M. Benkhalifa, "Implicit links based web page representation for web page classification," in *Proc. 5th Int. Conf. Web Intell., Mining Semantics*, Jul. 2015, pp. 1–11.

[30] Y. Jing and S. Baluja, "Pagerank for product image search," in *Proc. 17th Int. Conf. World Wide Web*, Apr. 2008, pp. 307–316.

[31] W. Zhou, Q. Tian, and H. Li, "Visual block link analysis for image re-ranking," in *Proc. 1st Int. Conf. Internet Multimedia Comput. Service*, Nov. 2009, pp. 10–16.

[32] L. Xie, Q. Tian, W. Zhou, and B. Zhang, "Fast and accurate near-duplicate image search with affinity propagation on the ImageWeb," *Comput. Vis. Image Understand.*, vol. 124, pp. 31–41, Jul. 2014.

[33] W. Zhang, C.-W. Ngo, and X. Cao, "Hyperlink-aware object retrieval," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4186–4198, Sep. 2016.

[34] X. Wang, W. Zhou, Q. Tian, and H. Li, "Adaptively weighted graph fusion for image retrieval," in *Proc. Int. Conf. Internet Multimedia Comput. Service*, Aug. 2016, pp. 18–21.

[35] H. Aouadi, M. T. Khemakhem, and M. B. Jemaa, "Building contextual implicit links for image retrieval," in *Proc. 20th Int. Conf. Enterprise Inf. Syst.*, 2018, pp. 81–91.

[36] H. Aouadi, M. T. Khemakhem, and M. B. Jemaa, "Uncovering hidden links between images through their textual context," in *Proc. 20th Int. Conf., Enterprise Inf. Syst. (ICEIS)*. Madeira, Portugal: Springer, Jul. 2019, pp. 370–395.

[37] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.

[38] T. Tsikrika, A. Popescu, and J. Kludas, "Overview of the Wikipedia image retrieval task at ImageCLEF 2011," in *Proc. CLEF Notebook Papers/Labs/Workshop*, vol. 4, 2011, p. 1–17.

[39] J. M. Kleinberg, "Authoritative sources in a hyperlinked environment," *J. ACM*, vol. 46, no. 5, pp. 604–632, Sep. 1999.

[40] S. Brin and L. Page, "The anatomy of a large-scale hypertextual Web search engine," *Comput. Netw. ISDN Syst.*, vol. 30, nos. 1–7, pp. 107–117, Apr. 1998.

[41] M. D. Dunlop, *Multimedia Information Retrieval*. Glasgow, U.K.: Univ. Glasgow, 1991.

[42] V. Harmandas, M. Sanderson, and M. D. Dunlop, "Image retrieval by hypertext links," in *Proc. 20th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, 1997, pp. 296–303.

[43] R. Lempel and A. Soffer, "PicASHOW: Pictorial authority search by hyperlinks on the web," in *Proc. 10th Int. Conf. World Wide Web*, Apr. 2001, pp. 438–448.

[44] E. G. M. Petrakis, "Intelligent search for image information on the web through text and link structure analysis," in *Multimodal Processing and Interaction Audio, Video, Text*. Berlin, Germany: Springer, 2008, pp. 1–17.

[45] E. Voutsakis, E. G. M. Petrakis, and E. Milios, "Weighted link analysis for logo and trademark image retrieval on the Web," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. (WI)*, Sep. 2005, pp. 581–585.

[46] D. Cai, X. He, W.-Y. Ma, J.-R. Wen, and H. Zhang, "Organizing WWW images based on the analysis of page layout and Web link structure," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, vol. 1, Jun. 2004, pp. 113–116.

[47] X. He, D. Cai, J.-R. Wen, W.-Y. Ma, and H.-J. Zhang, "Clustering and searching WWW images using link and page layout analysis," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 3, no. 2, p. 10, May 2007.

[48] D. Cai, S. Yu, J.-R. Wen, and W.-Y. Ma, "VIPS: A vision-based page segmentation algorithm," Tech. Rep., 2003.

[49] Y. Jing and S. Baluja, "VisualRank: Applying PageRank to large-scale image search," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 11, pp. 1877–1890, Nov. 2008.

[50] S. Zhang, M. Yang, T. Cour, K. Yu, and D. N. Metaxas, "Query specific fusion for image retrieval," in *Proc. 12th Eur. Conf. Comput. Vis. (ECCV)*. Florence, Italy: Springer, Oct. 2012, pp. 660–673.

[51] M. Wang, H. Li, D. Tao, K. Lu, and X. Wu, "Multimodal graph-based reranking for web image search," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4649–4661, Nov. 2012.

[52] Z. Liu, S. Wang, L. Zheng, and Q. Tian, "Robust ImageGraph: Rank-level feature fusion for image search," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3128–3141, Jul. 2017.

[53] R. Khasanova, X. Dong, and P. Frossard, "Multi-modal image retrieval with random walk on multi-layer graphs," in *Proc. IEEE Int. Symp. Multimedia (ISM)*, Dec. 2016, pp. 1–6.

[54] L. Chen, Y. Gao, Y. Zhang, S. Wang, and B. Zheng, "Scalable hypergraph-based image retrieval and tagging system," in *Proc. IEEE 34th Int. Conf. Data Eng. (ICDE)*, Apr. 2018, pp. 257–268.

[55] A. Barman and S. K. Shah, "A graph-based approach for making consensus-based decisions in image search and person re-identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 3, pp. 753–765, Mar. 2021.

[56] S. Liu, M. Sun, L. Feng, H. Qiao, S. Chen, and Y. Liu, "Social neighborhood graph and multigraph fusion ranking for multifeature image retrieval," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 3, pp. 1389–1399, Mar. 2021.

[57] L. P. Valem and D. C. G. Pedronette, "Graph-based selective rank fusion for unsupervised image retrieval," *Pattern Recognit. Lett.*, vol. 135, pp. 82–89, Jul. 2020.

[58] J. Liu, W. Lai, X.-S. Hua, Y. Huang, and S. Li, "Video search re-ranking via multi-graph propagation," in *Proc. 15th ACM Int. Conf. Multimedia*, Sep. 2007, pp. 208–217.

[59] W. H. Hsu, L. S. Kennedy, and S.-F. Chang, "Video search reranking through random walk over document-level context graph," in *Proc. 15th ACM Int. Conf. Multimedia*, Sep. 2007, pp. 971–980.

[60] M. Eskevich, G. J. F. Jones, R. Aly, R. J. F. Ordelman, S. Chen, D. Nadeem, C. Guinaudeau, G. Gravier, P. Sébillot, T. de Nies, P. Debevere, R. Van de Walle, P. Galuscakova, P. Pecina, and M. Larson, "Multimedia information seeking through search and hyperlinking," in *Proc. 3rd ACM Conf. Int. Conf. Multimedia Retr.*, Apr. 2013, pp. 287–294.

[61] S. Chen, M. Eskevich, G. J. Jones, and N. E. O'Connor, "An investigation into feature effectiveness for multimedia hyperlinking," in *Proc. 20th Anniversary Int. Conf., MultiMedia Model. (MMM)*. Dublin, Ireland: Springer, Jan. 2014, pp. 251–262.

[62] C. Bhatt, N. Pappas, M. Habibi, and A. Popescu-Belis, "Multimodal reranking of content-based recommendations for hyperlinking video snippets," in *Proc. Int. Conf. Multimedia Retr.*, Apr. 2014, pp. 225–232.

[63] A.-R. Şimon, R. Sicre, R. Bois, G. Gravier, and P. Sébillot, "IRISA at TRECVid2015: Leveraging multimodal LDA for video hyperlinking," in *Proc. TRECVid Workshop*, 2015, pp. 1–9.

[64] R. Bois, V. VukotiĆ, A.-R. Simon, R. Sicre, C. Raymond, P. Sébillot, and G. Gravier, "Exploiting multimodality in video hyperlinking to improve target diversity," in *Proc. 23rd Int. Conf., MultiMedia Model. (MMM)*. Reykjavik, Iceland: Springer, Jan. 2017, pp. 185–197.

[65] H. Aouadi, M. T. Khemakhem, and M. B. Jemaa, "An LDA topic model adaptation for context-based image retrieval," in *Proc. 16th Int. Conf. Electron. Commerce Web Technol., E-Commerce Web Technol. (EC-Web)*. Valencia, Spain: Springer, Sep. 2015, pp. 69–80.

[66] L. C. Freeman, "Centrality in social networks: Conceptual clarification," in *Social Network: Critical Concepts in Sociology*, vol. 1. London, U.K.: Routledge, 2002, pp. 238–263.

[67] U. Brandes, "A faster algorithm for betweenness centrality," *J. Math. Sociol.*, vol. 25, no. 2, pp. 163–177, 2001.

[68] A. Baccini, S. Déjean, L. Lafage, and J. Mothe, "How many performance measures to evaluate information retrieval systems?" *Knowl. Inf. Syst.*, vol. 30, no. 3, pp. 693–713, Mar. 2012.

[69] F. Wilcoxon, "Individual comparisons by ranking methods," in *Breakthroughs in Statistics*. Berlin, Germany: Springer, 1992.

[70] D. Hull, "Using statistical testing in the evaluation of retrieval experiments," in *Proc. 16th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, 1993, pp. 329–338.

[71] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Commun. ACM*, vol. 18, no. 11, pp. 613–620, Nov. 1975.

**KARIM GASMI** received the bachelor's degree in computer science, in 2008, and the master's and Ph.D. degrees in computer science from the University of Sfax. He is currently an Assistant Professor with Jouf University, Saudi Arabia. He is also a member of the Research Laboratory on Development and Control of Distributed Applications (Redcad). His research interests include information retrieval and medical image processing.

**HATEM AOUADI** received the master's and Ph.D. degrees in computer systems engineering from the National School of Engineers of Sfax, in 2010 and 2018, respectively. He is currently an Assistant Professor with the National School of Computer Sciences (ENSI), University of Sfax. His research interest includes image retrieval from the web, focusing on analyzing contextual information, such as links, topics, and metadata, to enhance retrieval results.

**MOUNA TORJMEN** received the Ph.D. degree in computer science from the Paul Sabatier University of Toulouse, France, in 2009. In September 2010, she was appointed as an Associate Professor with the Department of Computer Science, National School of Engineers of Sfax (ENIS), University of Sfax, and she joined the Research Laboratory on Development and Control of Distributed Applications. She is currently a Professor of ENIS, University of Sfax. More precisely, she is involved in building and testing relevant models to transform data into knowledge and values. Her research interests include information systems and data science. More details are available on her home page: http://redcad.org/members/mouna.torjmen/.