*International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*
**Web Site: www.ijettcs.org Email: editor@ijettcs.org**
**Volume 5, Issue 2, March-April 2016** **ISSN 2278-6856**

# Web Image Re-Ranking With Feedback Mechanism

**Rutuja N. Patil, Prof.Sandeep B.Vanjale**

[1]M.tech Research Scholar Dept. of Computer Engineering
Bharati Vidyapeeth Deemed University College of Engineering, Pune.

[2]Phd. Research Scholar Dept.Computer Engineering
Bharati Vidyapeeth Deemed University College of Engineering, Pune.

## Abstract

***Image Search Applications often uses key-Terms as Query and likewise is contingent on adjacent text to locate image. These search applications causes problem scenario as of vagueness of query Term as it's hard for lay-mans to appropriately represent visual content of objective image by merely expending terms. Re-ranking Image is an effectual method to progress consequences of web search applications this method is been applied by market giant search engines that as Google and yahoo. A foremost trial in investigation of re-ranking image is, resemblances of visual attributes not always link well with correct meanings of image that understand client's goal. Specified a query term a list of image is primarily found grounded on textual evidence. By probing user to select a probe picture from pool lasting images are again listed founded on their graphical ties with query picture. The attributes of images are anticipated into their interrelated correct spaces to get specific image. The query-exact semantic signs broadly expand both correctness and competence of Re-ranking image. This Research work we Implement Two systems offline search using Simrank algorithm and online search with API of commercial search Engine like Google Bing to retrieve Image clusters . One click Explicit feedback is been used to select appropriate cluster and after one click re-structured search results are been displayed. Performance of system is been evaluated on Precision and recall finding lesser time delay in online system but better object similarity in offline. Research effort demonstrate performance of two algorithms on time delay overall system is been tested on user search feedback and time delay and memory consumption.***

**Keywords:** Simrank, Image search, Image Retrieval System (IRS), Bing, word net, Information Retrieval.

## 1. INTRODUCTION

Today we largely we search information in form image as its better and quick way to understand information. Social networks like instagram flicker have been daily used by millions of user and are large source of knowledge generating network of Information. Information as is when
required related to particular subject can be retrieved by querying this systems..
the method reviewed in this paper, a query keyword is first used to retrieve a set of images based on the keyword. Then the user is asked to pick an image from these images. Also, the rest of the images are ranked based
on their visual similarities. The major challenge is the correlation of similarities of visual features and images' semantic meaning, which are needed to interpret users' intention to search. Recently, it has been proposed to match images in a semantic space that used attributes or reference classes closely related to the semantic meanings of images as basis. However, characterizing the highly diverse images from the web is difficult because it is impossible to learn a universal visual semantic space

## 2. BACKGROUND KNOWLEDGE

Background Knowledge related to filed to image retrieval or Picture Retrieval from large set. Image retrieval scheme is computer Machine for surfing, finding and retrieving pictures from huge catalogue of digital pictures. Maximum customary and common techniques of image retrieval apply particular technique of tallying metadata that as tags terms, and definitions related to image. Tag process has been most effective method in image retrieval which helps find relevant information from in precise.Manual Tag generation is time-overwhelming arduous and costly to overcome this great quantity of is on current state on image tag generation in automated format. Moreover surge in social online applications and web of knowledge have enthused expansion of numerous web-grounded image tag generation implementations.**[1].**In wide-ranging IRS queries can be classified as[1].

**[1.] Feature-based query:** This Method has technique where we find 'an' image by number

**Query:** Find 'an' OR Find image 12 June**.**

**[2]Textual query: Image retrieval Method that uses text as associated to retrieve images.**

**Query:** Find pictures "gateway of India".
Find pictures "Modi speech".

**[3]Visual query:** query technique based on attributes like color, texture, size and shape.

**Query:** Find pictures "red apple".
Find pictures "blue sky".
Find pictures "3.5, 4.5, red".

Query type's exploit various image DS (descriptors) and

necessitate diverse dispensation utilities. Image descriptors are categorized as:

**[1] Metadata DS (descriptors):**

**a.) Feature-grounded framework** and operational metadata that as creator period category image name, size,type.etc.

**b.) Text-grounded semantic** metadata that as headings/caption topic/terms lists free-text portrayals or text nearby entrenched images which are used in HTML.

**[2]Visual DS (descriptors):** extracted from picture through storage procedure by an picture retrieval machine as suggested and castoff by image reading civic.

**DS include:**

- ❖ High/pixel level attributes telling color feel or shapes within image.
- ❖ Object set recognized within an image.

Visual DS are cast-off to practice to generate image autographs which could be indexed. An image query is examined by means of the same DS technique giving a query sign that is then linked to image signature to regulate resemblance between query requirement and DB image sign.

**[2.]Content-based image retrieval:** query by image contented and content-founded visual data retrieval is application of computer prophecy methods to image retrieval problematic which is delinquent of probing for pictures in large databanks (CBIR) retrieval is different to outdated concept-founded methods [as compared to Concept-founded image indexing]

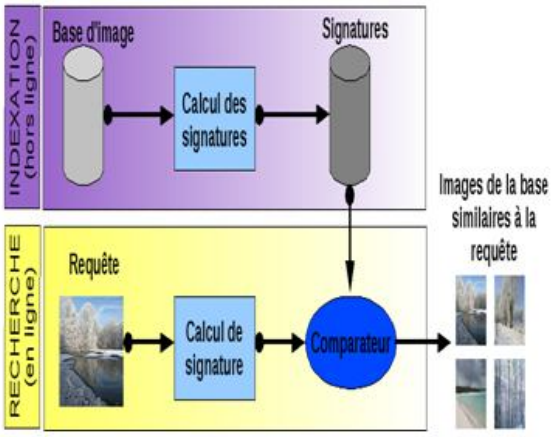


**Fig1:** CBIR System [1]

**These are following two techniques used in CBIR System**

**[1]Content DS color and texture [2] Recognizing shapes of image objects**

## 3 IMAGE RE-RANKING AND FEEDBACK

### 3.1 Image Re-ranking

Image re ranking, as an effective way to improve the results of web based image search, has been adopted by current commercial search engines. Given a query keyword, a pool of images is first retrieved by the search engine based on textual information. By asking the user to select a query image from the pool, the remaining images are re ranked based on their visual similarities with the query image. A major challenge is that the similarities of visual features do not well correlate with images' semantic meanings which interpret users' search intention. On the other hand, learning a universal visual semantic space to characterize highly diverse images from the web is difficult and inefficient

### 3.2 Feedback Mechanism

Feedback Mechanism is an alternative optimization technique used in to overcome algorithms limitations and increase search relevance with human intervention. feedback are classified as implicit and Explicit . Here in this work we propose Explicit Feedback. Implicit feedback would be added Future work .

### 3.3 Existing Approaches

In Existing system, one way is text based keyword expansion, making the textual description of the query more detailed. Existing linguistically related methods find either synonyms or other linguistic related words from thesaurus, or find words frequently co-occurring with the query keywords. In existing system low level visual features of images compared for re ranking purpose. For comparing visual features of images it uses Global weighting and adaptive weighting approaches. For example, Google image search provides the "Related Searches" feature to suggest likely keyword expansions. However, even with the same query keywords, the intention of users can be highly diverse and cannot be accurately captured by these expansions

Web image search engines use keywords as queries and search images based on the text associated with them. It is difficult for users to accurately describe the visual content of target images only using keywords and hence text based image search suffers from the ambiguity of query keywords. For example, using Apple as a query keyword, the retrieved images belong to different categories, such as apple laptop, apple logo, apple fruit.

To capture users' search intention, additional information has to be used in order to solve the ambiguity. Text based keyword expansion is one way to make the textual description of the query more detailed. Existing methods find either synonyms or other linguistic related words from thesaurus.

However, the intention of users can be highly diverse and cannot be accurately captured by these expansions, even with the same query keywords. Content based

image retrieval with relevance feedback is widely used in order to solve this ambiguity. Users are required to select multiple relevant and irrelevant image examples and the visual similarity metrics are learned through online training from them. Images are re ranked based on the learned visual similarities. However, for web scale commercial systems, users' feedback has to be limited to the minimum without online training

computing the visual similarities that reflect the

semantic relevance of images is the key component of image re ranking. Many visual features have been developed in recent years. However, the effective low level visual features are different for different query images

## 4. TABULATED LITERATURE SURVEY

| Author | Abstract | Methodology | Limitation and Future Scope |
|---|---|---|---|
| Rui | CBIR system has two limitations high level concepts and low features, with subjective human perception. This framework takes in user search at query perception subject is hold with weight of feedbacks. Relevance feedback is found to increase system performance and tested on 70000 images. | ❖ **Multimedia object model**<br>❖ **Integrating RF in CBIR: Get weights**<br>❖ **Find user search distribution**<br>❖ **Feature set ,calculate object similarity**<br>❖ **Normalization.**<br>❖ **Weight update.** | ❖ Various media format as audio video has to handle as image.<br>❖ Future scope is optimal Sub weight updating strategy.<br>❖ Expectation Maximization EM |
| Rasiwasia | QBVE (query by example ) and SR(semantic retrieval) fusion approach is been used .pictures are tagged with concept of dictionary .image is presented as vector. QBSE products RS which are large accurate than what was formerly conceivable. straight judgment of visual & semantic presentations under a mutual query model | ❖ Semantic Labeling System<br>❖ Semantic Multinomial<br>❖ Similarity Function | ❖ Advantage in designing retrieval systems with huge concept Classifications.<br>❖ multi-resolution semantic spaces |
| Jingyu | Most daily use search engines rely heavily on image and text to adjacent it for information retrieval with nosily solutions. text based query is re-ranked with adaptive visual resemblance | ❖ categorize image in predefined categories and definite feature to find similarity | ❖ photo quality re-ranking method is future work |
| Xin | Both link and | ❖ **H-Mok SimRank** | ❖ System has not been |

| | | | |
|---|---|---|---|
| | attributes similarities are been used for better Image Retrieval process. Image enriched network and performing recommendation is been done with SimRank. Fusion Approach is used to find image based on content and link association. | ❖ **2 Level algorithm**<br>❖ Algo1:[Link-based similarity measures].<br>❖ Algo2:[Content-based measure]<br>❖ IWSL achieves content and link bolstering Approach learning with whichever global or local attributes mass learning. | implemented on Distributed platform.<br>❖ Network partitioning is Accountable Future work.<br>❖ Time Complexity of Simrank is more and reducing this is challenge.<br>❖ Duplicate Image need to removed. |
| Ning | Images arranged and stored in to most related terms .cluster of image are generated on visual content. Approach is reducing uncertainatity in phrases and term associated to image. Relevance achieved as such high | ❖ Automated [image, text] couple generation.<br>❖ Instructive-image extraction.<br>❖ Related-text identification<br>❖ Text deconstructing and break apart with ranking.<br>❖ Automatic duo couple arrangement.<br>❖ Graph founded network of pictures and word are generated. | ❖ Need better web data Extraction system.<br>❖ Indexing reduces search time and hence needs large attention on better indexing. |
| Xiaogang | One click Feedback mechanism is used for re-ranking images in user preference of intent accounted from one click. Offline system trains on visual attributes with keyword expansion generating visual signs. Online level system work this signs reducing space to 25 attributes only and better performance. | ❖ Keyword-Enlargement and Mapping visual attributes to semantic space in group of semantic signs for keywords with query image choice.<br>❖ **Approach: Re-ranking**<br>❖ **Reference class selection**<br>❖ Single classifier for sign.<br>❖ Multiple classifiers for more than two sign. | ❖ keyword expansions Castoff to describe reference classes could be combined with other metadata and log files further textual and visual features.<br>❖ Incremental learning.<br>❖ Hashing is effective technique and needs to implement. |
| Shanmin | At situation's words in query are irrelevant and to get relevant works from query is an | ❖ **process1**→Initialize similar images to I.<br>❖ **process2**→Query visual word selection for scoring.<br>❖ dataset (oxford, Paris) | Memory consumption and lesser cost are issues to be handled. |

| | | | |
|---|---|---|---|
| | alternative optimization. A repetitive process is been done to align correct words with images. | | |

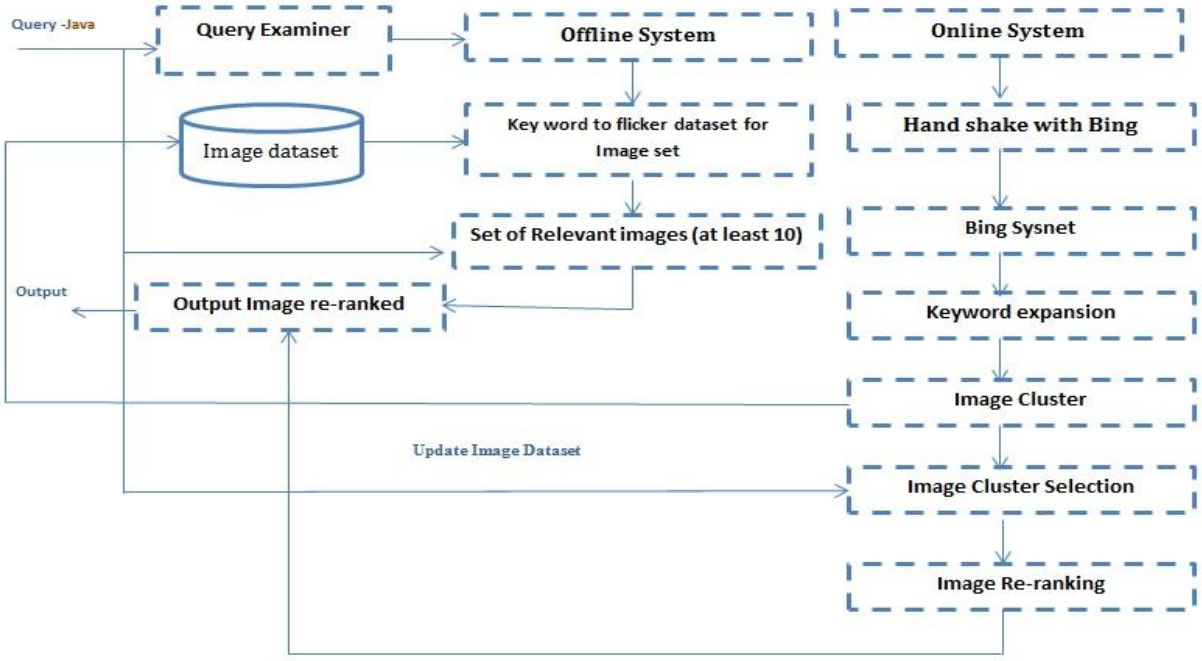


**Fig1: Architecture of proposed System**

**4.1Future Scope**.

**Scope [3]:** dictionary vocabulary is need for better results [wordnet is best scope]

**Scope [7]:** incremental learning and duplicate removal.

**Scope [8]:** formation of rich cross image and text dataset for <word, image > disambiguation.

**Scope [9]:** focused little cost counting both memory and computation.

**Scope [10]:** Work on Developing distributed-Computing with Product Recommendation System is Future and Foremost Scope of Work.

**4.2Problem Statement**

Keep it simple(KIS) based problem statement is been developed where research work is distributed in group of Task and incremental problem solving approach is been used.

Research Task 1: Develop and Design Dataset based Image Search on flicker Image Dataset implementing SIMRANK algorithm.

**Supervised Algorithm**

**SIMRANK Algorithm**

**Input:** G, the image-rich information network [generated from flicker dataset]
**Generate Tag schema ();**//170 tags
**Create Graph ();**//jung graph time required is ms 110//
**ReadGraph_file();**// serialized read by simrank is 27518 size 1098304.
**Basic features ();// sample image 1//**
**Process: Query Based Search**

1. Search Query: "Apple"(f).
2. Find top K similar candidates of each object;
3. **Firstservelt1 inilizied ().**

```
Void doGet(HttpServlet Request,HttpServletResponse)
throw Servlet Exception
{String my=(String)
getServletContext().getAttribute("MyVal");
System.out.println(my);
MyDataBasemydata=(MyDataBase)getServletContext().
getAttribute("Database");
Connection con=((MyDataBase) mydata).getCon();
response.setContentType("application/json");
```

❖**Research Work 1**

```
try {
String term = request.getParameter("term");
System.out.println("Data from ajax call " + term);
TagsLoader TagsLoaderObj = new TagsLoader();// Tag
name Required Apple//
ArrayList<String>TagList =
TagsLoaderObj.getAllTags(term,con);String searchList
= new Gson().toJson(TagList);
response.getWriter().write(searchList);
```

1. Iterate f

Tag List ==[tag 167]

Apple tag is available in graph

2. Compute link similarity for all image pairs;

```
List<String> getimageListForTagKey(String tagId,
Connection con) {
List<String> imageListForTag=new
ArrayList<String>();
Statement stmt=null;
```

Supervised Algorithm

SIMRANK Algorithm

**Research Task 2:** Develop and Design Image search Engine with WorldNet for Better Keyword set.

Research Task 3: Develop and Design Dataset based Image Search with complete handshake() with BING search Engine and Retrieve information from web in group of cluster (to reduce search time) and develop one click search for re-ranking

## 5. PROPOSED WORK

Proposed work is outcome of three research tasks RT's with Three Different Algorithms which are nothing but incremental upgaradation in procedure and better system development. The System has been Design as in Fig 2 with object oriented Design approach

```
try{
stmt=(Statement)
con.createStatement(ResultSet.TYPE_SCROLL_INSENSITIV
E, ResultSet.CONCUR_READ_ONLY);
String sql="Select distinct img_key from ImgTagSchema
where tag_key = '" + tagId + "'";
ResultSet rs=stmt.executeQuery(sql);
                    while (rs.next()){

imageListForTag.add(rs.getString(1));
}
System. out. Println ("Total images obtained from
getimage List For Tag Key() ="+image List For Tag.size())//
Compute link similarity for all group pairs;//
                    return imageListForTag;
//Compute link similarity for all tag pairs//
Map<String, Float> myResultSimRankMap = new
HashMap<String, Float>();
Map for fix node= up8={0.8,im900=0.546778}
Optimized Mapping ();
```

8. Until converge or stop criteria satisfied;
9. Perform feature learning;
10. Update image similarities;

**Output:** S, pair-wise node similarity scores. User selects a image and only relevant results are shown to user.

**Research Work 2**

Clustering Algorithm

**Clustering Algorithm**

Input: G, the image-rich information network [generated from flicker dataset]
Process: Online web search
1. Search Query: "Apple"
2. Send word net search

```
Get Word Net (String sreachKey) throws Malformed URL
Exception, IOException {
String wordForm = sreachKey; //Get the synsets
containing the word form=capicity
System. out .println("IN get Word Net::");
File f = new File("C:/dict");
System. set Property("wordnet.database.dir", f. to String());
DireNounSynset[] hyponyms;NounSynset nounSynset;
Word Net Database data base =
WordNetDatabase.getFileInstance();
Synset[] synsets = database.getSynsets(wordForm);
synsets retrieved nounSynset = (NounSynset) (synsets[0]);
hyponym nounSynset.getHyponyms();
System.out.println("hyponyms::"+hyponyms[1]) if
(synsets.length > 0) { ArrayList<String> al = new
ArrayList<String>// add elements to al, including
duplicates
```

3. Send The set of words retrived from word net to Bing
4. Iterate f for all words [w1, w2, w3….];
5. Compute link to Bing search engine for all images;
6. Compute link for all groups;
7. Create Clusters for link similarity for resulting images;
8. each cluster contains similar results.
9. Update image similarities;

Output: S, Cluster-wise similarity scores.
Output: S, pair-wise node similarity scores.

**Research Work 3**

Clustering Algo [Semi-Supervised]

**Clustering Algo [Semi-Supervised]**

**Input:** G, the image-rich information network [generated from flicker dataset]
**Process:**

1. Search Query: "Jakarta"
2. Send text query t;
   SELECT wordnet of direct option
3. Keyword expansion by Bing sysnet() for keyword "Jakarta"
   get Keywords();// INBUILT FUNCTIONS//
   W=[Jakarta=Indonesia-beaches,Jakarta-worldmap,Jakarta skyline,Jakarta airport ,Jakarta food,jakarta landmarks……..]
4. Iterate f for all W;
5. Compute link to Bing search engine for all images;

```
downloadallImages()// function//
BufferedImage image = null;
try {
if (!(imgSrc.startsWith("http"))) {
                    url = url + imgSrc;
                } else {
                    url = imgSrc;
                } imgSrc =
imgSrc.substring(imgSrc.lastIndexOf("/"
) + 1); String imageFormat = null;
                imageFormat =
imgSrc.substring(imgSrc.lastIndexOf("."
) + 1); imgSrc = imgSrc + "";
                System.out.println("imgSrc::::::::"
+ imgSrc);
```

**6.** Compute link for all groups;
**7.** Create Clusters for link similarity for resulting images;
Clusters=[Jakarta Indonesia,java,apple,…….]
**8.** Each cluster contains similar results;
9. User selects cluster and gives one click feedback
10. Re-ranked results of images only from relevant clusters are displayed.
11. Put in Tag() process .

**Output:** S, Cluster-wise similarity scores.

## 5. Evaluation of Work

**Table 1:** Research Results

| **Parameter** | *Value* | **Algorithm1** | **Algorithm2** | **Algorithm3** |
|---|---|---|---|---|
| keywords | **java** | No image | No image | Cluster:[Java island java programming, java |
| | **Jakarta** | No image | No image | Jakarta place, Jakarta island |
| | **Apple** | 3 images | 23 images | [Apple ,apple company, red apple, green apple] |
| | **Virat kholi** | No image | 32 images | Virat anushka, virat sachin,virat kholi |
| | | | | |
| | **Mickey** | **No image** | **7 Image** | **Mickey mouse, red mickey,** |
| Time Delay | **apple** | **10234ms** | **2345ms** | **2954ms** |
| | **sports** | **10567ms** | **987ms** | **456ms** |
| User Feedback | **[Good,better,fai,worst]** | **fair** | **better** | **Good** |

**Table2:** Time Delay

| **Parameter** | **Algorithm1** | **Algorithm3** |
|---|---|---|
| **Time delay** | **5.9** | **1.5** |
| **Search overhead** | **492** | **135** |
| | | |

**Table3:** Feedback Evaluation for 18 key queries

| Parameter | Algorithm1 | Algorithm3 |
|---|---|---|
| **Best** | **2** | **10** |
| **Better** | **5** | **5** |
| **Good** | **11** | **3** |
| **Worst** | **0** | **0** |

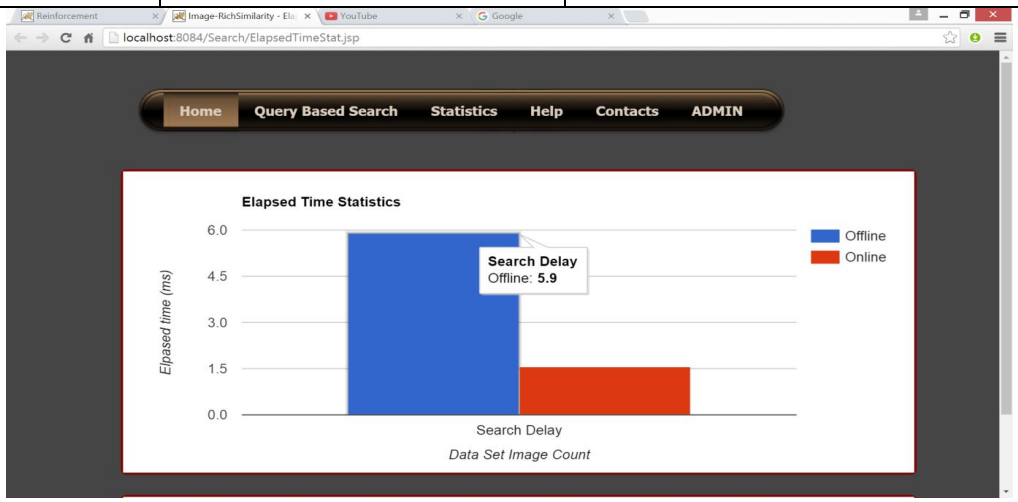


**Fig 3:** Research snapshot 1

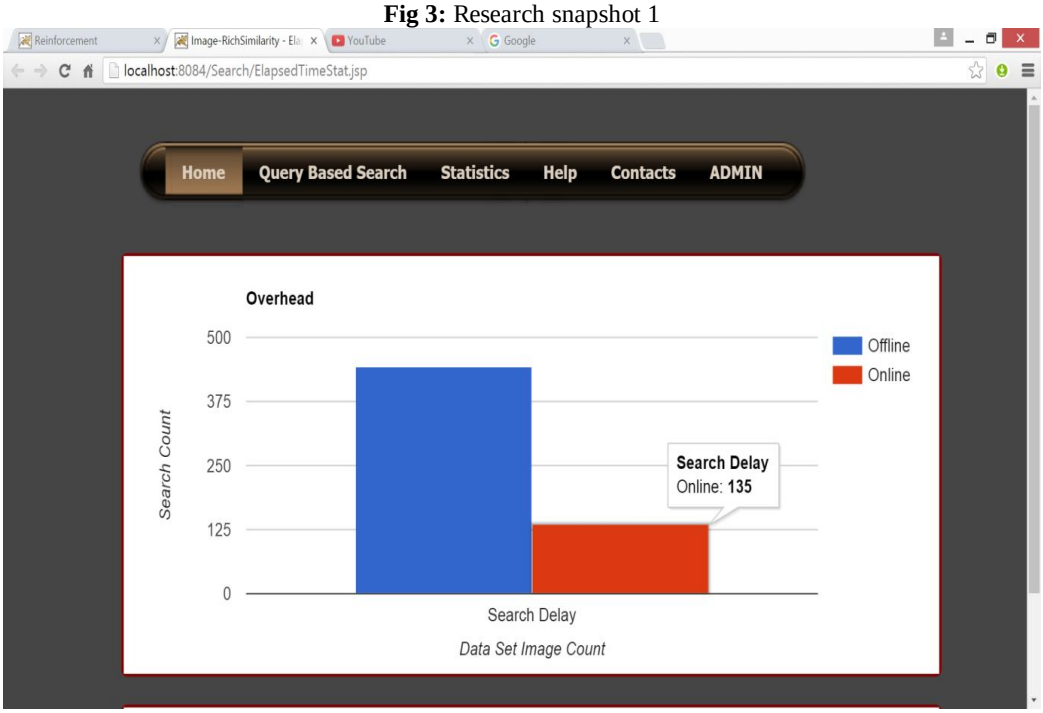


**Fig 4:** Research snapshot 2

## 6. CONCLUSION

The proposed system has been implemented for image and can be extended to video audio search .Currently our system demonstrate comparative on online offline system but we can combine this procedure to form system which detects if information not present then only query online else find it in offline system and also need to work on different evaluation parameters.

## References

[1] www.wikipdeia.org [ online source].
[2] Y. Rui, T. S. Huang, M. Ortega, et.al, "Relevance feedback: a power tool for interactive Content-based image retrieval", IEEE Transactions On Circuits and Systems for Video Technology, 1998.
[3] N. Rasiwasia, P. J. Moreno, et.al, "Bridging the gap: Query by semantic example", IEEE Transactions. On Multimedia, vol. 9, no. 5, pages.923 -938, August 2007.
[4] J.Cui, F. Wen, et.al, "Real time Google and live image search re-ranking", The 16th ACM international conference on Multimedia, Pages 729-732, 2008.
[5] XinmeiTian, Dacheng Tao et.al, "Active Re-ranking for Web Image Search",IEEE Transactions on Image Processing, Vol. 19, No. 3,March 2010.
[6] X. Tang, K.Liu, J. Cui, et.a, "IntentSearch: Capturing User Intention for One-Click Internet Image Search", IEEE Transactions On Pattern Analysis and Machine Intelligence Vol. 34, No.7 pages 1342 – 1353, July 2012.
[7] XinJin,JieboLuo,Jie Yu et. al,"Reinforce Similarity Integration In Image Rich Information Network", IEEE Transactions on Knowledge & Data Engineering ,vol.25,Issue No.02, Feb 2013.
[8] Xiaogang Wang, Ke Liu et.al,"Web Image Re-Ranking Using Query-Specific Semantic Signatures", IEEE Transactions on Pattern Analysis and Machine Intelligence Volume: 36 , Issue: 4April 2014.
[9] Ning Zhou, Jianping Fan, "Automatic image–text alignment for large-scale web image indexing and retrieval", Pattern Recognition Elsevier July 2014, 05–219.
[10] Shanmin Pang, Jianru Xue, ZhanningGao, QiTian, "Image re-ranking with an alternating optimization", Neuro-computing, Elsevier, march 2015.

## AUTHORS

**Rutuja N. Patil** is currently pursuing M.Tech (Computer) from Department of Computer Engineering, Bharati Vidyapeeth Deemed University College of engineering Pune, India. She received her B.Tech (Computer) Degree from Bharati Vidyapeeth Deemed University College of engineering Pune, India. Her area of interest include data mining, web mining.

**Prof. Sandeep Vanjale** is working as a Professor in computer engineering department at Bharati Vidyapeeth University College of engineering, Pune, Maharashtra, India. He received his ME (Computer) degree from Bharati Vidyapeeth University College of engineering, Pune. His research interests include Computer Network, Network Security, WLAN Security. He attended more than 40 national and international conferences and published 50 papers in international conference and Journals.