A Novel Web Image Re-Ranking Approach Based on Query Specific Semantic Signatures

Ms. Pooja P. Dutta
ME Student
Department of Computer Engineering
Narnarayan Shastri Institute of Technology, Jetalpur, Ahmedabad, GTU-Gujarat, India

Asst. Prof. Anand Chauhan
Assistant Professor
Department of Mathematics
Narnarayan Shastri Institute of Technology, Jetalpur, Ahmedabad, GTU-Gujarat, India

Abstract

Image re-ranking, is an effective way to improve the results of web-based image search. Given a query keyword, a pool of images are initailly retrieved primarily based on textual data, the remaining images are re-ranked based on their visual similarities with the query image corresponding to the user input. A major challenge is that the similarities of visual features don't well correlate with images' semantic meanings that interpret users' search intention. Recently people proposed to match pictures in a semantic space that used attributes or reference categories closely associated with the semantic meanings of images as basis. Even though, learning a universal visual semantic space to characterize extremely diverse images from the internet is troublesome and inefficient. In this thesis, we propose a completely distinctive image re-ranking framework that learns completely different semantic spaces for numerous query keywords automatically at the on-line stage. The visual features of images are projected into their corresponding semantic spaces to induce semantic signatures. At the online stage, images are re-ranked by scrutiny their semantic signatures obtained from the semantic spaces such that by the query keyword. The proposed query-specific semantic signatures considerably improve both the accuracy and efficiency of image re-ranking.
Keywords: Web Image, Web Image Re – Ranking, Image Query, Content Based Retrieval, Semantic Signatures, Visual Similarities, Keywords Expansion

---

I. INTRODUCTION

Web - scale image search engines principally use keywords, surrounding text to search pictures, and content based image retrieval with relevance feedback. But, since the internet is a large collection of images as compared to web-scale systems, the retrieval of images accurately, according to the user's intention becomes a major problem. Many web image search engines have adopted the 'Online image re-ranking' strategy, which limits users' effort to merely one-click feedback.

Web image search engines mostly use keywords as queries and search pictures supporting the text linked to them. It's a little troublesome for users to accurately describe the visual content of target pictures needed, hence only mistreatment keywords and thus text-based image search suffers from the anomaly of question keywords. For instance, mistreatment of apple as a question keyword, the retrieved pictures belong to fully different categories, like apple laptop computer, apple logo, apple fruit. To capture users' search intention, extra data needs to be used so as to resolve the anomaly. Keyword expansion based on text is a technique to create the textual description of the query in additional details. Existing strategies search either synonyms or totally different linguistic-related words from reference book. But, the intention of users may be extremely different and cannot be accurately captured by these expansions, even with equivalent question keywords.

Content-based image retrieval with relative feedback is widely utilised in order to resolve this ambiguity. Users are needed to decide out the multiple relevant and irrelevant image examples and additionally the visual similarity metrics are learned through on-line training from them. Images are re-ranked according to the learned visual similarities. However, for web-scale business systems, users' feedback needs to be restricted to the minimum while not on-line training. [2]

The recently proposed technique is a one in which when the user inputs a query, a pool of images are retrieved that are having similar visualities with that to the query input. The user is asked to select a pool that is having image similar to his search query image, and then the pictures are re ranked according to the semantic space of the images in the pool with the semantic space of the question input [1]. But, characterizing the highly various images from the internet is troublesome because it's not possible to learn a universal visual semantic space. [2]

II. RELATED WORK

A huge development has been made within the area relating to search image from the web corresponding to the user input. The various standard ways like content based retrieval, attribute dominance [3], pseudo relevance [15], mapped visual features to a universal concept dictionary for image retrieval [5] that did not use query images.

The key part of image re-ranking is to find visual similarities presenting linguistics relevancy of images. Many visual choices [8] are developed in recent years. But, for various query images, the effective low-level visual options are completely different.

Cui et al. [9] classified query pictures into eight predefined intention categories and gave completely different feature weighting schemes to differing types of query images. Yet, it's tough for the eight coefficient schemes to hide the huge diversity of all internet images. It's also possible for a question image to be classified to a wrong class. So as to cut back the linguistics gap, query-specific linguistics signature was initially planned in [11]. Kuo et al. [12] recently augmented each image with relevant linguistics choices through propagation over a visible graph and a matter graph that were correlate.

Later the conventional framework was noted that used query images so as to create the actual search of image consistent with the user's intention. Then this method too had an improvement wherever the semantic space of the query input of the user was compared with the semantic signatures of the pictures within the search pool retrieved[1].

The various methods for the web image re ranking are as follows.

A. *Intent Search [6]:*

In this technique, a completely distinctive web image search approach is represented. It solely needs the user to click on one query image with minimum effort and pictures from a pool retrieved by text-based search square measure re ranked supporting each visual and TEXT content. The key contribution is to capture the users' search intention from this one-click question image in four steps.

1) The question image is categorized into one in each of the predefined adaptative weight categories that replicate users' search intention at a rough level. Within each category, a specific weight schema is utilized to combine visual options adaptative to the present fairly image to higher re rank the text-based search result.
2) Supported the visual content of the question image designated by the user and via image clump, question keywords area unit distended to capture user intention.
3) Distended keywords area unit accustomed enlarge the image pool to contain extra relevant images.
4) Distended keywords also are accustomed to expand the question image to multiple positive visual examples from that new question specific visual similarities and textual similarity metrics area unit learned to additional improve content-based image reranking of these steps are at unit automatic, while not extra effort from the user.

B. *Attribute Dominance [3]:*

When we explore an image, some properties or attributes of the image stand out quite others. Once describing an image, people are most likely to explain these dominant attributes first. Attribute dominance could be a results of a fancy interplay between the various properties present or absent within the image, which attributes in a image are additionally dominant than others and reveals rich data regarding the content of the image. This technique, emphasis on data, by modeling attributes dominance. It tends to show that this helps improve the performance of vision systems on a variety of human-centric applications like zero-shot learning, image search and generating matter descriptions of pictures.

The method consists of four steps as follows.

1) Annotating Attribute Dominance:

Here images are annotated in attribute dominance train the attribute dominance predictor. The dominance annotations are collected at the class level, although approach trivially generalizes to image level dominance annotations as well.

2) Modeling Attribute Dominance:

Given a novel image x_t, we predict the dominance $d^m{}_t$ of attribute m in that image using

$$d^m{}_t = w^T{}_m\varphi(x_t)$$

We represent image xt via an image descriptor. We use the output scores of binary attribute classifiers to explain the image. This exposes the complex interplay among attributes mentioned in the introduction that results in the dominance of sure attributes in a picture and not others. The relevant aspects of the interplay ar learnt by our model. φ(xt) can be simply disturbance or an implicit high- (potentially infinite-) dimensional feature map implied by a kernel. For training, we project the category-level attribute dominance annotations to each training image.

3) Zero-shot Learning:

In zero-shot learning, the supervisor describes novel N' previously unseen classes in terms of their attribute signatures, n' ∈. With a pre-trained set of M binary classifiers for each attribute Direct Attribute Prediction (DAP) model, the probability that an image x belongs to each of the novel categories Cn' is

$$pa_{n'}(x)\ \alpha\ {}^{M}\pi_m{=}1pa^m(x)$$

Where, $pa^m(x)$ is the probability that attribute a^m takes the value $g^m{}_{n'} \in \{0, 1\}$ in image x as computed using the binary classifier for attribute a^m. The image is assigned to the category with the highest probability $pa_{n'}(x)$. This approach forms our baseline. It relies on an interface where a supervisor goes through every attribute in a pre-defined arbitrary order and indicates its presence or absence in a test category. [3]

4) Image Search:

Here, the image search situation wherever a user has a target category in mind is taken into account, and provides as question an inventory of attributes that describe that category. It's unlikely that the user can provide the values of all M attributes once describing the question. (S)he is probably going to use the attributes dominant within the target construct, naming the foremost dominant attributes first.

5) Automatic Query Expansion [4]:

This technique explores the ways to derive higher object models given the query space, so as to enhance retrieval performance. It tends to maintain the form of the model fixed: it's still a configuration of visual words. However, instead of merely extracting the model from the one input question region, it tends to enrich it with extra data from the corpus; and tend to examine with latent model represented during this technique.

C. An overview of the Approach is as follows [4]:

1) Given a query space, search the corpus and retrieve a set of image regions that match the query space. We use bag-of-visual-words retrieval on with special verification, but the approach would apply to retrieval systems that use completely different object models.
2) Mix the retrieved regions, together with the first query, to create a richer latent model of the interest criteria of items.
3) Re-query the corpus exploitation this swollen model to retrieve a swollen set of matching regions.
4) Repeat the method as necessary, alternating between model refinement and re-querying.

D. Query by Semantic Example [5]:

A combination of query-by-visual-example (QBVE) and linguistics retrieval (SR), denoted as query-by-semantic-example (QBSE), is used during this technique. Pictures are labeled with regard to a vocabulary of visual concepts, as is common in SR. Every image is then depicted by a vector, cited as a linguistics multinomial, of posterior concept possibilities. Retrieval relies on the question-by-example paradigm: the user provides a query image, for which 1) a linguistics multinomial is computed and 2) matched to those among the data. QBSE is shown to own two main properties of interest, one mostly sensible and additionally the alternative philosophical.

From a practical point of view, because it inherits the generalization ability of SR among the space of acquainted visual concepts (referred to because the linguistics space) however performs much better outside of it, QBSE produces retrieval systems that are extra correct than what was previously potential. Philosophically, as a result of it permits a direct comparison of visual and linguistics representations beneath a standard question paradigm, QBSE permits the planning of experiments that expressly check the value of linguistics representations for image retrieval. An implementation of QBSE to a lower place the minimum chance of error (MPE) retrieval framework, antecedently applied with success to every QBVE and SR, is planned, and used to demonstrate the two properties. Specifically, an in depth objective comparison of QBSE with QBVE is given, showing that the previous considerably outperforms the latter each among and outside the linguistics space. By rigorously dominant the structure of the linguistics space, it's additionally shown that this improvement will solely be attributed to the linguistics nature of the illustration on that QBSE depends.[5]

E. Conventional Image Re – ranking Framework [1]:

Online image re-ranking that limits users' effort to merely one-click feedback is an efficient means to improve search results and its interaction is straightforward enough. Major internet image search engines have adopted this strategy. Its diagram is shown in Fig. 1. Given a query keyword input by a user, a pool of pictures relevant to the question keyword area unit retrieved by the computer program corresponding to a stored word-image index file. Typically the size of the came back image pool is mounted, e.g., containing 1,000 pictures. By asking the user to pick out a query image, which demonstrates the user's search intention, from the pool, the remaining pictures among the pool space unit re-ranked depending on their visual similarities with the question image. The word image index file and visual options of images area unit pre-computed offline and stored.1 the most on-line method value is on comparison visual features. To achieve high efficiency, the visual feature vectors got to be short and their matching has to be quick. Some standard visual features space unit in high dimensions and efficiency is not satisfactory if they are directly matched. [1]

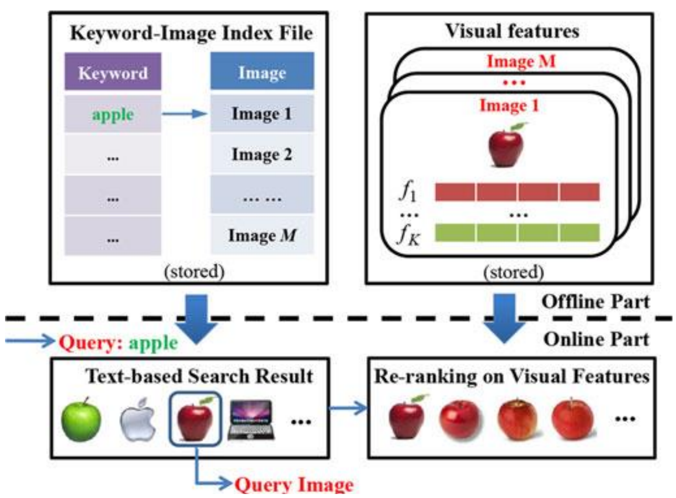

Fig. 1: Conventional Image Re – Ranking Framework

III. EXISTING SYSTEM

Another major challenge while not on-line training, is that, the similarities of low-level visual features may not well correlate with the images' high-level linguistics meanings that interpret users' search intention. Moreover, low-level options are sometimes not consistent with perception. As an example, if pictures of a similar object are captured from completely different viewpoints, beneath completely lightings or maybe with different compression artifacts, their low-level options might change significantly, though humans suppose the visual content does not change a lot of, to decrease this semantic gap and inconsistency with perception, there are selection of studies to map visual choices to a group of predefined ideas or attributes as semantic signatures [7], as an example, Kovashka et al. [7] planned a system that refined the image search with relative attribute feedback. Users described their search intention with reference images and a collection of pre-defined attributes. These ideas and attributes are pre-trained offline and have tolerance with variation of visual content. However, these approaches are exclusively applicable to closed image sets of comparatively small sizes, however not appropriate for on-line web-scale image re-ranking.

Since the topics of net images amendment dynamically, it is fascinating that the ideas and attributes are typically automatically found instead of being manually outlined.

In this method, a completely unique framework is planned for web image re-ranking. Rather than manually shaping a universal idea dictionary, it learns totally different linguistics areas for various query keywords severally and automatically. The semantic space associated with the images to be re-ranked is considerably narrowed down by the question keyword provided by the user, for instance, if the query keyword is "apple," the ideas of "mountain" and "Paris" space unit irrelevant and should be excluded. Instead, the ideas of "computer" and "fruit" are going to be used as dimensions to search out out the linguistics house related to "apple." The query-specific semantic areas can more accurately model the images to be re-ranked, since they have excluded different probably unlimited variety of irrelevant concepts that serve solely as noise and deteriorate the re-ranking performance on every accuracy and computing worth. The visual and matter features of images area unit then projected into their connected semantic areas to urge semantic signatures. At the online stage, pictures space unit re-ranked by scrutiny their linguistics signatures obtained from the semantic space of the question keyword. The linguistics correlation between ideas is explored and incorporated once computing the similarity of linguistics signatures.

Experiments show that the semantic area of a question keyword is delineating by merely 20-30 ideas (also referred as "reference classes"). Thus, the semantic areas are terribly short and on-line image re-ranking becomes very economical. As a result of the large type of keywords and also the dynamic variations of the net, the linguistics spaces of question keywords space unit automatically learned through keyword growth.[1]

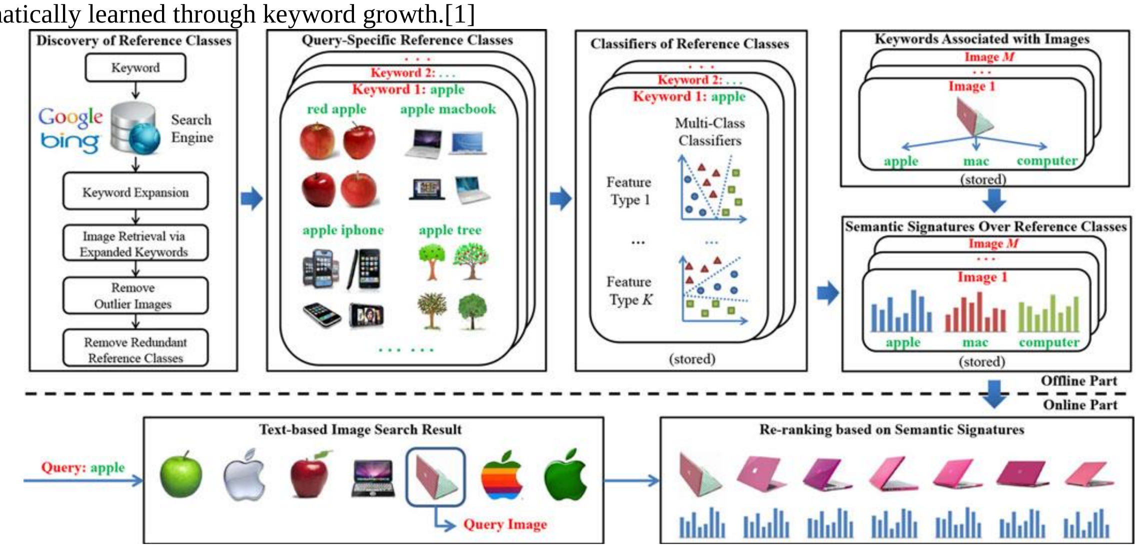

Fig. 2: Image Re – Ranking Framework using Semantic Signatures

IV. PROPOSED SYSTEM

Here a novel approach is proposed to re rank the searched images based on the semantic signatures of the images. The semantic signatures of the images are generated previously in the offline stage and stored and when the imagequery is made the images are re ranked based on the semantic signatures of these images retrieved based on the keywords. The modules of the system are as follows:

A. Re-Ranking Accuracy:

In this module, we include five labelers to manually label testing images under each query keyword into different categories based on semantic meanings. Image classes were carefully defined by the five labelers through inspecting all the testing images under a query keyword. Defining image categories was completely independent of discovering reference classes. The labelers are unaware of what reference classes have been discovered by our system. The number of image categories is also different than that of the number of reference classes. Each image was labeled by at least three labellers and its label was decided by voting. Some images irrelevant to query keywords were labeled as outliers and not assigned to any category.

B. Re-Ranking Images outside Reference Class:

If the category of a query image corresponds to a reference class, we deliberately delete this reference class and use the remaining reference classes to train classifiers and to compute semantic signatures when comparing this query image with other images.

C. Incorporating Semantic Correlations:

We can then after incorporate semantic correlations between reference classes when computing image similarities. For each type of semantic signatures obtained above, we compute the image similarity. The re-ranking precisions for all types of semantic signatures on the three data sets. Notably, the data sets achieves around 10 percent relative improvement compared with other ones, reaching the performance multiple despite its signature is six times shorter.

D. Re-Ranking with Semantic Based:

Query-specific semantic signature can also be applied to image re-ranking without selecting query images. This application also needs the user input as query keyword. But it assumes that images returned by initial text-only search have a dominant topic and images that are belonging to that topic should have higher ranks. Our query-specific semantic signature is quite effective in this application since it can improve the similarity measurement of images.

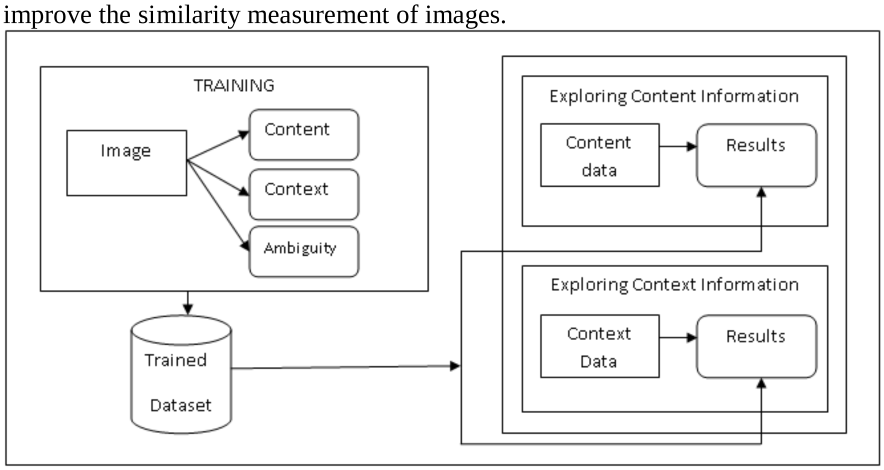

Fig. 3: Block Diagram of the Proposed System

V. CONCLUSION

Web Image Re ranking is one among the popular research areas, where a huge number of techniques are proposed for it. Starting with the Content based Retrieval to semantic signatures a immense work has been done, where the images are retrieved and re ranked based on attributes, keywords, visual features, their dominant attributes, relevance feedback of the user, pseudo relevance feedback and many more. But the re ranking with semantic signatures has been quite effective in this space, since it learns the semantic signatures offline and then uses these semantic signatures for image retrieval at the online stage thereby consuming the time only for the online stage. This new technique overcomes the existing system in terms of efficiency, diversity of images retrieved, reducing space of semantic signatures and storage dimensions

REFERENCES

[1] Xiaogang Wang, Shi Qiu, Ke Liu, and Xiaoou Tang. "Web Image Re-Ranking UsingQuery-Specific Semantic Signatures." IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 36, No. 4, April 2014
[2] Harshil Jain, Riya Lodha and Khushali Deulkar. "Web-based Image Search using Image Re-Ranking Technique: A Review Paper" International Journal of Current Engineering and Technology, Vol.4, No.5, Oct 2014
[3] Naman Turakhia and Devi Parikh. "Attribute Dominance: What Pops Out?"IEEE International Conference on Computer Vision 2013
[4] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman, "Total Recall: Automatic Query Expansion with a Generative Feature Model for Object Retrieval," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2007.
[5] N. Rasiwasia, P.J. Moreno, and N. Vasconcelos, "Bridging the Gap: Query by Semantic Example," IEEE Trans. Multimedia, vol. 9, no. 5, pp. 923-938, Aug. 2007.
[6] Xiaoou Tang, Ke Liu, Jingyu Cui, Fang Wen and Xiaogang Wang. "IntentSearch: Capturing User Intention for One-Click Internet Image Search". IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 34, NO. 7, JULY 2012
[7] A. Kovashka, D. Parikh, and K. Grauman, "WhittleSearch: Image Search with Relative Attribute Feedback," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2012.
[8] J. Philbin, M. Isard, J. Sivic, and A. Zisserman, "Descriptor Learning for Efficient Retrieval," Proc. 11th European Conf. Computer Vision ECCV), 2010.
[9] J. Cui, F. Wen, and X. Tang, "Real Time Google and Live Image Search Re-Ranking," Proc. 16th ACM Int'l Conf. Multimedia, 2008.
[10] J. Cui, F. Wen, and X. Tang, "Intent Search: Interactive on-Line Image Search Re-Ranking," Proc. 16th ACM Int'l Conf. Multimedia, 2008.

[11] X. Wang, K. Liu, and X. Tang, "Query-Specific Visual Semantic Spaces for Web Image Re-Ranking," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2010.
[12] Y. Kuo, W. Cheng, H. Lin, and W. Hsu, "Unsupervised Semantic Feature Discovery for Image Object Retrieval and Tag Refinement," IEEE Trans. Multimedia, vol. 14, no. 4, pp. 1079-1090, Aug. 2012.
[13] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-Based Image Retrieval," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.
[14] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-Based Image Retrieval," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.
[15] R. Yan, E. Hauptmann, and R. Jin, "Multimedia Search with Pseudo-Relevance Feedback," Proc. Int'l Conf. Image and Video Retrieval, 2003.
[16] R. Datta, D. Joshi, and J.Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age," ACM Computing Surveys, vol. 40, article 5, 2007.
[17] Y. Rui, T.S. Huang, M. Ortega, and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," IEEE Trans. Circuits and Systems for Video Technology, vol. 8, no. 5, pp. 644-655, Sept. 1998.
[18] D. Tao, X. Tang, X. Li, and X. Wu, "Asymmetric Bagging and Random Subspace for Support Vector Machines-Based Relevance Feedback in Image Retrieval," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 28, no. 7, pp. 1088-1099, July 2006.
[19] J. Krapac, M. Allan, J. Verbeek, and F. Jurie, "Improving Web Image Search Results Using Query-Relative Classifiers," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2010.