TOWARDS MORE EFFICIENT FOR WEB IMAGE SEARCH ENGINE USING DOMINANT MEANING TECHNIQUE

Yasser Ibrahim[1], Mohammed Abdel Razek[1, 2], Kamal A. ElDahshan[1]

[1]Math. & Computer Science Department, Faculty of Science, Azhar university, Nasr City,11884,Cairo,Egypt.
[2]Deanship of Distance Learning, King Abdulaziz University Kingdom of Saudi Arabia, P.O. Box 80254.
yasser.ibrahim@azhar.edu.eg, maabdulraze1@kau.edu.sa, dahshan@gmail.com

KEY WORDS: Image Retrieval, Web Search Engine, Dominant Meaning Technique.

ABSTRACT:

A huge type of media Online is creating big challenges for researcher to retrieve a right media for a wright query. This paper describes a suggested formwork to improve retrieving images from a Web search engine using a dominant meaning method. The dominant meaning is a set of keywords that best fit an intended meaning of a target query. This technique sees the target meaning as a master word and the words that clarify it as slaves. The hypnosis of this research is the more accurate the slaves, the more accurate the retrieving results. The proposed re-ranking algorithm is built based on the dominant meaning. The contribution of the research is located on the construction of the query and then the filtering of incoming results from the Web search. The research proposes a calculation method of the rank values for Web search engine results. The proposed procedures starting with constructing a considerable query from its slaves, sending a query to Our Custom Search Image Engine (as Web Image Search Engine), Retrieve images, retrieving each images document, extracting dominant meaning words from each document included image, calculating rank values of each document based on dominant meaning probability for each document, re-sorting retrieved images based on the ranking value, and then posting sorted images to a user. To ensure the proposed system could run successfully, we conducted an empirical analysis on a real example of examination.

1. INTRODUCTION

The search by using original query usually retrieve huge numbers of results which made noisy results and waste time of users. We found that the statistical analysis more users are interested in top one hundred results so needed to find new technique to filter the results. Let us do the next exercise, this exercise will use Google Image Search Engine (as Web Image Search Engine) and try search by using queries Q1 is "Faculty of science", Q2 is "Faculty of science al azhar", Q3 is "Faculty of science al azhar cairo".

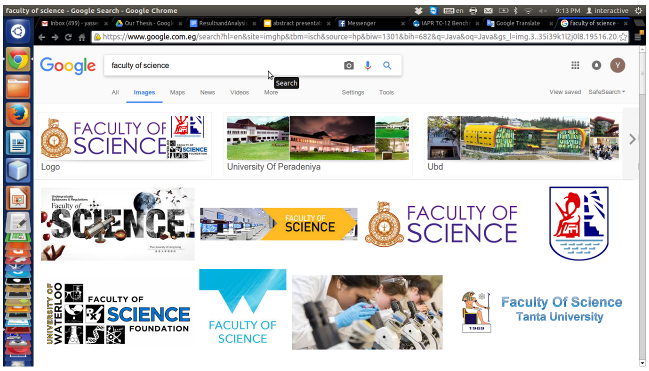

Figure 1: Retrieved Results From Google Image Search Engine for query Q1 ("Faculty of science")

Figure 1. shows the retrieved results from Google Image Search Engine which contains all images relevant by faculty of science and university in any region in the world, but Figure 2. shows filtration of the retrieved results which contains all images relevant by faculty of science in egypt this filtration caused by adding al azhar word to Q1 and Figure 3 shows more filtration of the retrieved results than Q2, cased by adding cairo word to Q2.

From this exercise we conclude that the length of query is an important factor to reduce retrieved results. This conclusion raises

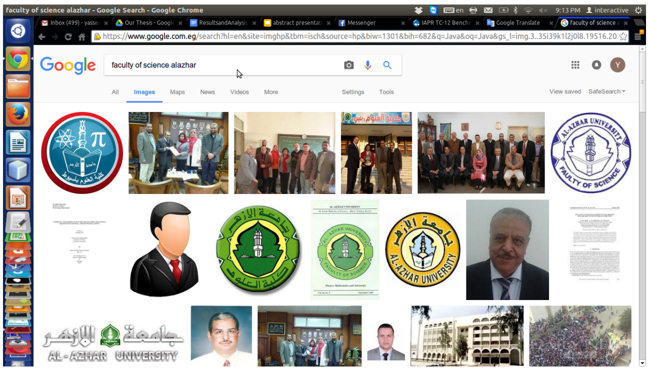

Figure 2: Retrieved Results From Google Image Search Engine for query Q2 ("Faculty of science al azhar")

a question What the best length of the query? and this is what we will try to answer.

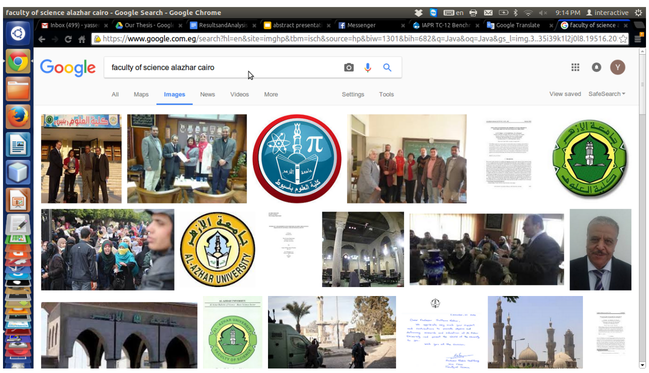

Figure 3: Retrieved Results From Google Image Search Engine for query Q3 ("Faculty of science al azhar cairo")

We have organized the rest of this paper as follows. Section 2. briefly describes some related work; Section 3. shows how the proposed method; Section 4. presents an overview of our proposed architecture; Section 5. presents some experimental results; plus Section 6.2 suggests future projects.

2. RELATED WORK

Labeled Image datasets is an important area to understanding image, so according to efforts in this area, generally these efforts divided into three categories:

1. Manual annotation.
2. Active learning.
3. An automatic methods.

Manual annotation was important way to construct image dataset in early years (e.g., STL-10 (Coates et al., 2011), CIFAR-10 (Deng et al., 2009), PASCAL VOC 2007 (Everingham et al., 2010), ImageNet (Deng et al., 2009) and Caltech-101 (Griffin et al., 2007)). This method introduced a high accuracy, but is labor intensive. To overcome cost of manual annotation method the researchers was focused on active learning method. Li et al. (Collins et al., 2008) introduced randomly labeled. Siddiquie et al. (Siddiquie and Gupta, 2010) presented an active learning framework to simultaneously learn contextual models for scene understanding tasks (multi-class classification). Grauman et al. (Vijayanarasimhan and Grauman, 2014) presented an approach for on-line learning of object detectors, but manual annotation and active learning depends on pre-exist annotations, which often results in one of the most significant limitations to construct a large scale image dataset. Automatic annotation is reduced the cost of manual annotation too, Schroff et al. (Schroff et al., 2011) adopted text information to rank images retrieved from a web search and used these top-ranked images to learn visual models to re-rank images once again. Li et al. [2] leveraged the first few images returned from an image search engine to train the image classifier, Yao et al. (Yao et al., 2016) proposed the use of multiple query expansions instead of a single query in the process of initial candidate images collection. Finally there are many efforts which not aimed to construct image dataset, but this efforts to generation query expansions and noisy image filtered, WordNet (Miller, 1995) and ConceptNet (Speer and Havasi, 2013) are often used to obtain synonyms to overcome the download restriction of these search engines.

Content based image retrieval uses visual feature to calculate image similarity. The method used to learn visual similarity metrics to capture users search intention was Relevance feedback. But it required more effort of users to select multiple relevant , irrelevant image examples and further needs online training. Cui et al. (Tang et al., 2012) proposed image re-ranking approach which limited users effort to just one click. This easiest image re-ranking approach method is adopted by many web scale image search engines like Google and Bing. Now a days there are a numbers of efforts on using predefined concepts or attributes as image signature for general image recognition and matching. Rasiwasia et al. (Rasiwasia et al., 2007) mentioned effect of visual features to a universal concept dictionary. Lampert et al. (Lampert et al., 2009) used predicted attributes to detect novel object classes with semantic meanings. Measuring the similarities between novel object classes and known object classes few approaches transferred knowledge between object classes. Relevance feedback techniques was started in late 60's and transferred to image domain as CBIR Systems.

3. PROPOSED METHOD

The query terms often contains of insignificant terms or terms need to clarification, in another meaning the query terms need to add or remove some words to increase intended meaning from the query. Dominant meaning approche one from approach which working increase meaning of query, and also length of the query terms from major factors to enhance retrieved results.
In this thesis, we will try overcome the noisy images which produce from query expansions by trying to answer about how many words which added to original query to improve retrieved results? To overcome these challenges, we will use dominant meaning approach (Razek, 2013). The dominant meaning approach is a set of keywords that best fit an intended meaning of a target word (Razek, 2013). This technique sees a query as a target meaning plus some words that fall within the range of that meaning and freezes up the target meaning, which is called a master word, and adds or removes some slave words, which clarify the target meaning razek2013towards, [(Razek and Kaltenbach, 2006), to build hierarchy meaning model and reconstruct new query from original query and use it in search engine instead of original query to minimize scope. In this thesis, we used Apache Lucene lucenewebsite which is open-source search software to build our proposed algorithm to search the Web based on a custom search engine (CSE), by indexing dataset and send query to find matched images, we show it in detail in chapter three.
The following steps are representing suggested approach:

1. Collect dataset(Metadata about classified images) for specific domain.
2. Build dominant meaning hierarchy.
3. Generate expansion queries of size 5, 10, , 45.
4. Design a proposed algorithm.
5. Develop the proposed algorithm to search the Web.
6. Run the proposed algorithm to execute the expansion queries with size ranged from 5, 10, ,into 45
7. Compute Precision, Recall and F1-Measure for all queries.
8. Compare the results and write the discussion.

4. PROPOSED SYSTEM OVERVIEW

In Figure 4, Architecture of proposed methodology. In this model we firstly Indexing Dataset to construct Indexing Storage, the indexing process is very important in our model, from this process there are two processes will start from it. The indexing storage we will use it to generate dominant meaning hierarchy and generate expanded queries next construct Dominant Meaning Hierarchical by applying Dominant Meaning algorithm. And then construct a list of Expanded Queries from original query by using Dominant Meaning Hierarchical. The of length 1 called original query or traditional query, Dominant Meaning Hierarchical can provides us list of slaves words which are try clarify the target meaning for traditional query we generate from list of slaves words list of queries of length 5, 10, , 45 which are represents list of expanded queries. Then we send traditional query and list of expanded queries to our Custom Search Engine, we build our Custom Search Engine by using apache Lucene (Apache, 2017) api which provides Java-based indexing and search technology, as well as spell checking, hit highlighting and advanced analysis/tokenization capabilities. Now retrieved results retrieved our Custom Search Engine, we will evaluate our methodology by compute precision, recall and F1 for each traditional query and list of expanded queries in an attempt to answer the question of the thesis. The following sections in this chapter we will show our architecture in details.

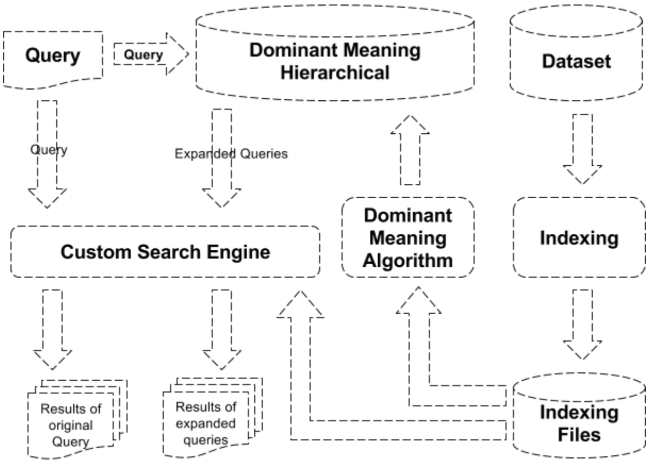

Figure 4: Architecture of proposed methodology

To constructing Dominant Meaning Hierarchical, suppose we have dataset consists of n, that $C = \{C_i\}_{i=1}^{n}$ is, for each C_i from C represented by a collection of documents which trying to describe concept C_i , suppose that the collection consists of m documents, that is $\{C_i = \{D_j^i, j = 1, 2, 3, ..., m_i\}\}$, each document in this collection consists of a set of k words or $D_j^i = \{w_{lj}^i\}, l = 1, 2, 3, ..., k_j\}$ terms. The w_{lj}^is represent the frequency of word w_{lj}^i occurs in document D_j which belongs to concept C_i. This frequency is computed as the number of times that the w_l occurs in the D_j^i. the following steps represent the process to choose top-N words which can represent the Dominant Meaning of concept C_i, suppose that word w_c^i symbolizes concept C_i.

- Calculate the values of $w_{lj}^i \forall i, j$.
- Suppose that $C_{i,j}$ is the frequency of concept C_i, which appears in document $D_j^i where j = 1, 2, 3, ..., m_i$.
- Calculate maximum of $C_{ji} \forall i, F_c^i = Max\{C_{ji}\}_{j=1}^{m_i}$.
- Calculate maximum value of $w_{lj}^i \forall l, j, F_{wj}^i = Max\{w_{lj}^i\}_{j=1}^{m_i}$.
- Choose P_c^i, which satisfies $0 F_{wj}^i F_c^i$.
- Finally, consider the dominant meaning probability: $P_{ij} = P_{ij}(w_j|C_i) = \frac{1}{m_i}[\sum_{j=1}^{m_i} \frac{w_{lj}^i}{F_c^i}], j = 1, 2, ..., k_j; i = 1, 2, ..., n$ (Dominant Meaning Formula).

Now for each concept C_i, we rank the terms of collection $\{P_{i1}, P_{i2}, ..., P_{im}\}$ in decreasing order. As a result, the Dominant Meaning of the concept C_i can be represented by the set of words that corresponds to the set $\{P_{i1}, P_{i2}, ..., P_{iN}\}$; that is $W_i = \{w_1^i, w_2^i, ..., w_N^i\}$. The set W_i is represent more intended meaning for concept C_i.

5. EXPERIMENTAL RESULTS

To try answer on question how many words which added to original query to improve retrieved results?. We conduct experiment on collection of classified images with Description.

6. DATA SET

The data set consist of 20,000 images, each image is associated with a text caption in up to three different languages (English, German and Spanish) IAPR TC-12 (CLEF, 2009). We already worked on 2719 images as training dataset, and too we label each category by using dominant meaning formula Razek (Razek et al., 2003). This formula is using probabilistic model. Table 1 shows describe the above training dataset.

Concept	Number of images
Background	616
Blue	674
Brown	430
Green	636
Gray	318
Snow	45
Total	**2719**

Table 1: Number of images in training dataset

6.1 Precision and Recall

In this research, the traditional keyword approach is represented when the query size is equal one. As we see the results show the purpose of using dominant meaning for a new query expansion method by which improves precision without degrading recall. Table 2 shows the results of Precision and Recall Results for the concept "Background" for different values for the query size. The query size is built by the dominant meaning method.
The initial result of the original query gives 61.54 and 40.27 for Precision and Recall respectively as shown in Table 2. The improvement in precision due to dominant meaning method is taken place around 0.5%. The best improvement is given at the query size contains 45 dominant meaning words, however, we found that the original query has better precision when the query size has 5 dominant meaning words. In contrast, the improvement in precision due to dominant meaning method is taken place around 48%.

Query Size	Precision	Recall
1	61.5385	40.2685
5	60.625	65.1007
10	61.75	82.8859
15	62.0192	86.5772
20	61.8056	89.5973
25	61.9048	91.6107
30	62.306	94.2953
35	62.4724	94.9664
40	62.5821	95.9732
45	62.7451	96.6443

Table 2: Precision and Recall Results for the concept "Background"

For recall, we can see that the best improvement happens when the query size contains 45 dominant meaning words and the lower one happened at the query size includes 5 dominant meaning words.

Figure 5 shows the precision of retrieval with our Dominant Meaning Approach (DMA) v.s. the traditional keyword based approach for the concept "Background".

As shown in Figure 5, the recall using dominant meaning words increases, while that the precision using dominant meaning words decreases as the number of retrieved documents increases in the interval between one word and 5 words. We can see that the recall is almost when using the dominant meaning terms than using

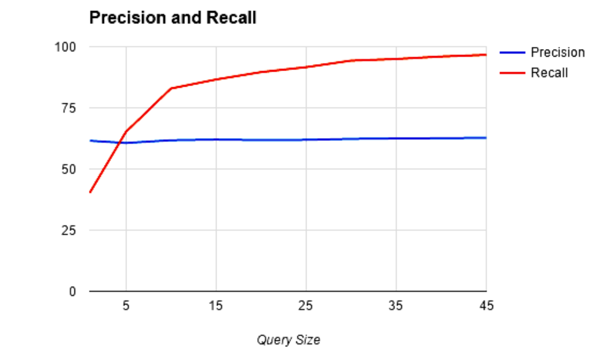

Figure 5: Dominant meaning approach v.s. the traditional keyword based approach for the concept "Background"

the original query, however, the precision is better when using number of words greater than 5 dominant meaning words.

Figure 6 F1- Measure for Precision and Recall Results for the concept "Background". It shows the results for five query expansion. The sequence of the queries contains 5, 15, 25, 35, ..., 45 terms respectively. On average for all recall/precision levels, F1-measure using a dominant meaning approach makes a good effect on the 5 levels of query expansion (5 to 45).

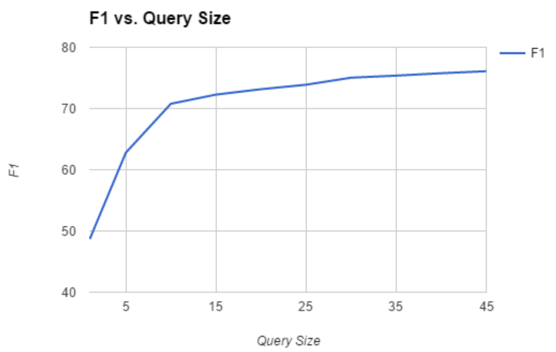

Figure 6: F1-Measure for Precision and Recall Results for the concept "Background"

Figure 7, shows the precision values for all concepts. At the size query is equal to five, Snow concept presents the best results and Background concept shows a lowest one.

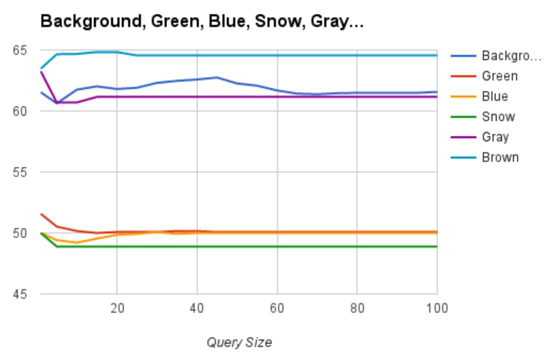

Figure 7: Precision Value for all Concepts

6.1.1 Precision

As shown in Table 3. The improvement in precision due to dominant meaning method for concepts Background, Snow and Brown is taken place around 1.31%, 1.11% and 1.09% respectively and for concepts Green and Gary, which they do not increase but a shortage by 1.49% and 2.1% respectively, while there has been no increase for concept Blue. It is noticeable that there are increased slightly for some concepts and don't occurred increase for another concepts. By comparing between the numbers, we got and the definition of precision this lead to the retrieved results from expanded queries contains many from irrelevant results, which affected in compute Precision. We found that the above the irrelevant results which retrieved from beside the relevant results because of there are similarity between dominant meaning hierarchy for concepts with each other. We know that this is a big issue and we will work on it in the future to solve this issue.

Concept	From	To	Improve...
Background	61.55%	62.75%	1.2066
Green	51.75%	50%	-1.4936
Blue	50.00%	50.00%	0
Snow	50.00%	49%	1.1111
Gray	63%	61%	-2.0992
Brown	63.47%	65%	1.0839

Table 3: The Improvement values in Precision

As shown in Figure 7, Table 4 Background and Brown concepts, the improvement in Precision is increased gradually for all the values of the query size and the precision was from 61.53% to 62.57%. For brown concept, the improvement in precision was from 63.47% to 64.56%.

6.1.2 Recall

Figure 8 represent Recall values for all concepts. At the size query is equal five all concepts are starting in increase, Snow concept presents the best results and Background concept shows a lowest one.

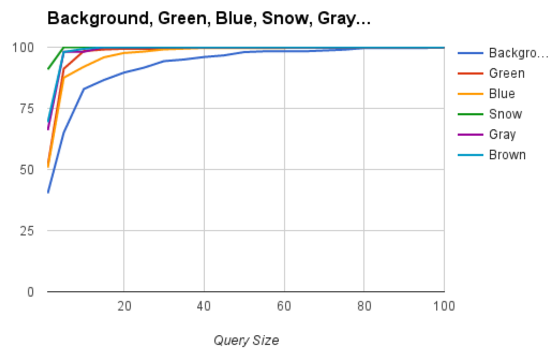

Figure 8: Recall value for all Concepts

Table 6 shows the improvement in recall due to dominant meaning method for concepts Background, Green, Snow, Gray and Brown is taken place around 56.3%, 48%, 49%, 9%, 34% and 31% respectively.
In general there are improvement clearly in recall values which tells us that our proposed method how well it has contributed in improvement image retrieved results.

As shown in Figure 8 and Table 5, All concepts, the improvement in Recall is increased quickly in the first values of the query size and be stable in remaining values. It is interesting to note that the highest improvement in recall obtained for concept Background,

Query Size	Background	Green	Blue	Snow	Gray	Brown
1	61.5385	51.5723	50	50	63.253	63.4731
5	60.625	50.5226	49.4137	48.8889	60.7004	64.6552
10	61.75	50.1603	49.2063	48.8889	60.7004	64.6808
15	62.0192	50	49.5399	48.8889	61.1538	64.8305
20	61.8056	50.0792	49.8485	48.8889	61.1538	64.8305
25	61.9048	50.0792	49.9246	48.8889	61.1538	64.557
30	62.306	50.0792	50.075	48.8889	61.1538	64.557
35	62.4724	50.1577	49.9255	48.8889	61.1538	64.557
40	62.5821	50.1577	50	48.8889	61.1538	64.557
45	62.7451	50.0787	50	48.8889	61.1538	64.557

Table 4: Represents all recall values for all concepts

Query Size	Background	Green	Blue	Snow	Gray	Brown
1	40.2685	51.5723	50.7418	90.9091	66.0377	69.281
5	65.1007	91.195	87.5371	100	98.1132	98.0392
10	82.8859	98.4277	91.9881	100	98.1132	99.3464
15	86.5772	99.0566	95.8457	100	100	100
20	89.5973	99.3711	97.6261	100	100	100
25	91.6107	99.3711	98.2196	100	100	100
30	94.2953	99.3711	99.1098	100	100	100
35	94.9664	100	99.4065	100	100	100
40	95.9732	100	99.7033	100	100	100
45	96.6443	100	99.7033	100	100	100

Table 5: Represents all precision values for all concepts

Concept	From	To	Improve...
Background	40.27%	96.60%	56.33
Green	51.60%	100%	48
Blue	50.70%	99.70%	49.00
Snow	90.90%	100%	9
Gray	66%	100%	34
Brown	69.34%	100%	31

Table 6: The Improvement values in Recall

the lowest one was for the concept Snow. The Recall value for Snow start from 90.9% for the original query and increased when the query size (5 words) with Recall value 100%.
From above discussion which lead to the our proposed method from numbers which we got for measurement values, the proposed method enhance the value of recall and there are enhance slight in precision values, which we will work on it in the future.

6.2 Conclusion & Future work

This research was trying answer the question how many words added to original query to improve retrieved results by using dominant meaning. To answer above question, we used dataset images with english description for each image, we apply dominant meaning approach to generate hierarchy meaning model and reconstruct new queries of different sizes $(5, 10, ...)$, and send to the original query and the new queries to our search engine and calculate precision, recall and F1-measure. The experimental results tell us the best improvement is given at the query size contains 45 dominant meaning words.

REFERENCES

Apache, 2017. Lucene. https://lucene.apache.org/.

CLEF, I., 2009. Dataset. http://imageclef.org/photodata.

Coates, A., Ng, A. and Lee, H., 2011. An analysis of single-layer networks in unsupervised feature learning. In: Proceedings of the fourteenth international conference on artificial intelligence and statistics, pp. 215–223.

Collins, B., Deng, J., Li, K. and Fei-Fei, L., 2008. Towards scalable dataset construction: An active learning approach. Computer Vision–ECCV 2008 pp. 86–98.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. In: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE, pp. 248–255.

Everingham, M., Van Gool, L., Williams, C. K., Winn, J. and Zisserman, A., 2010. The pascal visual object classes (voc) challenge. International journal of computer vision 88(2), pp. 303–338.

Griffin, G., Holub, A. and Perona, P., 2007. Caltech-256 object category dataset.

Lampert, C. H., Nickisch, H. and Harmeling, S., 2009. Learning to detect unseen object classes by between-class attribute transfer. In: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE, pp. 951–958.

Miller, G. A., 1995. Wordnet: a lexical database for english. Communications of the ACM 38(11), pp. 39–41.

Rasiwasia, N., Moreno, P. J. and Vasconcelos, N., 2007. Bridging the gap: Query by semantic example. IEEE Transactions on Multimedia 9(5), pp. 923–938.

Razek, M. A., 2013. Towards more efficient image web search. Intelligent Information Management 5(06), pp. 196.

Razek, M. A., Frasson, C. and Kaltenbach, M., 2003. Web course self-adaptation. In: Intelligent Agent Technology, 2003. IAT 2003. IEEE/WIC International Conference on, IEEE, pp. 614–617.

Razek, Mohammed Abdel, F. C. and Kaltenbach, M., 2006. Dominant meaning approach towards individualized web search for learning environments. Advances in web-based education: personalized learning environments p. 46.

Schroff, F., Criminisi, A. and Zisserman, A., 2011. Harvesting image databases from the web. IEEE transactions on pattern analysis and machine intelligence 33(4), pp. 754–766.

Siddiquie, B. and Gupta, A., 2010. Beyond active noun tagging: Modeling contextual interactions for multi-class active learning. In: Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE, pp. 2979–2986.

Speer, R. and Havasi, C., 2013. Conceptnet 5: A large semantic network for relational knowledge. In: The Peoples Web Meets NLP, Springer, pp. 161–176.

Tang, X., Liu, K., Cui, J., Wen, F. and Wang, X., 2012. Intentsearch: Capturing user intention for one-click internet image search. IEEE transactions on pattern analysis and machine intelligence 34(7), pp. 1342–1353.

Vijayanarasimhan, S. and Grauman, K., 2014. Large-scale live active learning: Training object detectors with crawled data and crowds. International Journal of Computer Vision 108(1-2), pp. 97–114.

Yao, Y., Zhang, J., Shen, F., Hua, X., Xu, J. and Tang, Z., 2016. Automatic image dataset construction with multiple textual metadata. In: Multimedia and Expo (ICME), 2016 IEEE International Conference on, IEEE, pp. 1–6.