International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 04 Issue: 08 | Aug -2017 www.irjet.net p-ISSN: 2395-0072

A Graph-based Web Image Annotation for Large Scale Image Retrieval

Khasim Syed[1], Dr. S V N Srinivasu[2]

[1]Research Scholar, Department of CSE, Rayalaseema University, Kurnool - 518002, (AP) - India
[2]Professor & Principal Department of CSE, IITS, Markapur

***

Abstract - Image annotation is a well-designed alternative to unambiguous recognition in images and an effective research topic in current years due to its potentially large impact on both image perceptive and Web image search. Using traditional methods, the annotation time for large collections is very high, while the annotation performance degrades with increase in number of keywords. In order to improve the accuracy of the image annotation, we proposed a framework called graph-based Web Image Annotation for Large Scale Image Retrieval. In this paper, our goal is clear the automatic image annotation issue in a new search and mining framework. First in searching stage, it identifies a set of visually similar images from a large-scale image database which are considered useful for labeling and retrieval. Then in the mining phase, a graph pattern matching algorithm is applied to find mainly representative keywords from the annotations of the retrieved image subset. To be able to rank relevant images, the approach is extended to Probabilistic Reverse Annotation. Our framework shows that the proposed algorithm improves effectively the image annotation performance.

***Key Words*: Reverse annotation, Web Image, Corel, Graph matching, visual concepts.**

1. INTRODUCTION

Automatic image annotation has received broad attentions in current years. The capability to search the content of text images is necessary for the usability and reputation. This is a challenge because: i) OCRs are not robust sufficient for web image mining ii) the text images hold large number of degradations and artifacts; iii) scalability to huge collections is tough. Moreover, users are expecting that search organization accept text queries and retrieve related outcome in interactive times. Since the arrival of economical imaging devices, the number of digital images and videos has full-grown exponentially. Huge collections of images and videos are at the present available and shared online also. Effective retrieval images from such big collections of multimedia figures are becoming an significant problem.

In the past years of image retrieval, images were annotated manually. Since manual effort was expensive and time taken process, this was reasonable for mostly military and medical domains. Content Based Image Retrieval (CBIR) systems have exposed adequate promise for querying-by-example, but the image matching methods are often computationally rigorous and thus time exhausting. Transcriptions of images through object recognition and scene analysis, etc. These methods not been effective, partly due to the limited applicability of present day's recognition techniques. Recently, recognition approaches have been confirmed for image retrieval, where a search index is built in a characteristic space. However, the trendy images or video retrieval systems are based on text, such as Google Images, which indexes multimedia with the nearby text. The fame is typically due to the interactive retrieval times for text based systems, as well as being capable to query-by-text. As a result, there has been a growing interest in automatic annotation of images. In this context, we proposed a graph-based Web Image Annotation for Large Scale Image Retrieval.

The rest of the paper is organized as follows. Before proceeding further, we shall look at existing annotation techniques and the issues to be addressed for building retrieval systems in the next section. We describe the Reverse annotation and our framework in Section 3. Implementation details are then reported in section 4. We conclude this paper and discuss some future works in section 5.

2. RELATED WORK

Several approaches have been proposed for annotating images by mining the web images with neighboring descriptions. A series of research were also done to influence information in world-wide web to annotate universal images. Given a query image, they first searched for related images from the web, and then mined delegate and frequent descriptions from the nearby descriptions of these related images as the annotation for the query image. C. V. Jawahar et al. [6] proposed a conversion model to label images at area level under the assumption that every blob in a visual vocabulary can be interpreted by certain word in a dictionary. Jeon et al. [4] proposed cross-media relevance model to expect the probability of generating a word given the blobs in an image. In the state that every word is treated as a distinct class, image annotation can be viewed as multi-class classification problem.

Wang et al. [5] proposed a search-based annotation scheme – AnnoSearch. This scheme requires a preliminary keyword as a seed to speed up the search by influence text-based search technologies. However, the early keyword might not for all time be available in real surroundings. In the case there is no early keyword available for the query image, the system will come across a serious efficiency problem. Furthermore, the system tends to be biased by the value of preliminary keywords. If the early keywords are not accurate, the annotation performance will disintegrate. Yang et al. [6] use multiple-instance learning to detect particular keywords from image data using labeled bags of examples. The fundamental intuition is to learn the most representative image region for a given keyword. Xirong Li et al. [2] uses content-based image retrieval (CBIR) facilitated by high-dimensional indexing to find a set of visually similar images from a large-scale image database. The database consists of images crawled from the World Wide Web with wealthy annotations. Based on search technologies, this framework does not force an explicit training phase, but efficiently leverages large-scale and well-annotated images, and is potentially capable of dealing with limitless vocabulary.

Towards the target of huge scale annotation, Yu Tang Guo et al [1] presented an approach called Reverse Annotation. Unlike time-honored annotation where keywords are identified for a given image, in Reverse Annotation, the significant images are identified for every keyword. With this ostensibly simple shift in perspective, the annotation time is reduced significantly. To be able to rank relevant images, the method is extended to Probabilistic Reverse Annotation. To advance the accuracy of the image annotation, Pramod Sankar K et al [4] proposed an automatic image annotation method based on mutual K-nearest neighbor graph (MKNN). The algorithm describes the association between low-level features, annotation words and image by a mutual K-nearest neighbor graph. Semantic information is extracted by exploiting the mutual relationship of two nodes in the mutual K-nearest neighbor graph. Inverse document frequency (IDF) is introduced to alter the weights of edges between the image node and its annotation word's node, which overcomes the deviation caused by high-frequency words.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

3. WEB IMAGE ANNOTATION SYSTEM

The intention of image annotation is to discover a keyword set L* that maximizes the conditional probability P (L | Eq), where w is a keyword in the vocabulary and Eq is an uncaptioned image, as indicated by (*) in Eq. 1. We reformulate this optimization problem from a search and mining perspective, that is,

$$L^* = \arg\max P\left(L \middle| E_q\right)$$

$$= \arg\max \sum_{I_i \in \varphi} P(L|E_i)\ P\left(IE_i \middle| E_q\right)$$

Where Ei denotes an image in a database φ, in which each image has some textual descriptions, P(Ei |Eq) denotes the probability that Ei is relevant (i.e. similar) to Eq, and P(L|Ei) represents the likelihood that Ei can be interpreted by L.

The issues toward large scale annotation are addressed by a novel Reverse Annotation framework. Reverse Annotation takes-off from the fact that, for any retrieval system, the number of items in the index is limited, while the images that are indexed could be unlimited. With a careful selection of the keywords/concepts for annotation, a large percentage of documents can be indexed, as well as a large number of user queries can be retrieved. In traditional annotation, the image features are compared with keyword models, where the annotation problem was that of classification. A multi class classifier or a hierarchy of classifiers is required to be learnt for this purpose.

Conversely, in Reverse Annotation, the keyword/label is matched with the images, and the matches are annotated with the keyword. This effectively is a verification problem, which involves only a two-class classification. Evidence from the biometrics community suggests that the verification problem has better performance than a classification problem. Thus, the annotation performance is expected to improve. The Reverse Annotation procedure has addressed the goal of identifying the relevant images for a given keyword. This is useful to build an index for search and retrieval. However, the indexed images cannot be ranked based on relevance, since the associations in the index are binary: either the image is relevant for a given keyword, or it is not. Ranking of the images is necessary to retrieve images that are more relevant to the given search query. The Reverse Annotation framework provides a mechanism for probabilistic annotation. For this purpose, we extend the framework to Probabilistic Reverse Annotation. It can be calculated as

$$p1_{il} = \frac{d(g_i|h_l)}{\sum_{l=1}^{m}(g_i|h_l)}$$

Where d(x//y) is a similarity measure (inverse of a distance measure). Within each cluster, the probability p2lj is estimated for the image region tlj to belong to tl. This probability can be estimated as

$$p2_{jl} = \frac{d\left(h_l \middle| h_{l_j}\right)}{\sum_{j=1}^{n_l^{\prime\prime}} d\left(h_l \middle| h_{l_j}\right)}$$

The total probability pij that cluster tlj belongs to keyword ki is given by

$$p_{ij} = p1_{il} * p2_{lj}$$

The aggregate weight of an image to a specified keyword is measured by accumulating the whole probabilities from every region of the image. We reveal the Reverse Annotation framework over the domain of text document images. There is a affluent variety of keywords in documents, which can be simply sampled from a text corpus. Keyword exemplars could be simply generated by translation the text to word images. The scalability of the approach could be tested over huge collections of images and it is comparatively simple to assess a retrieval system over text document images. The images from an image database are first preprocessed to advance their quality. These images then undergo various transformations and characteristic extraction to generate the significant features from the images. With the generated features, mining can be passed out using data mining techniques to discover significant patterns.

The resulting patterns are evaluated and interpreted to obtain the final knowledge, which can be applied to applications.

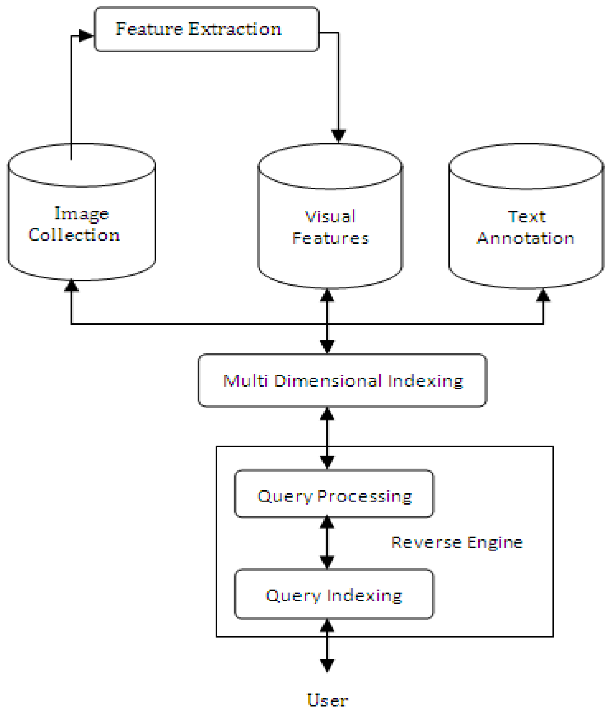


Fig-1 Large Scale Retrieval System for Web Image Mining

4. PREDICTION AND RANKING

Web document images ranking can also plays a main role. This is because throughout any search, if all the keywords in a user's query happen in a single slice of a page, the top ranking outcomes of pages returned by a search engine will be normally relevant and useful. However, if the query words spread at dissimilar segments in a page, the pages returned will not really always be relevant. If we can segment a page to its micro information units, then retrieval and ranking algorithms can utilize this additional information to return those most relevant pages. It first extracts salient phrases and measures a number of properties, such as expression frequencies, and combines the properties into a salience score based on a pre-learnt regression model. We propose the following control factor to make sure the visual consistency in Fs while avoiding and bringing in much noise:

$$\gamma = \frac{\text{Similarity(Bottom ranked Image, Query image)}}{\text{Similarity(Top ranked Image, Query image}}$$

γ Can be viewed as a tradeoff parameter between image quantity and quality, which is empirically set to 0.8 in this system.

5. METHODOLOGY

Imagine that a training set of annotated images is existing, and is represented as follows: D = $\{(I_i, x_i, y_i)\}$, i= 1.. N with x_i representing the feature vector of image Ii and yi denoting the annotation of that image. An annotated test set is also available and is represented by S = $\{(I_i, x_i, y_i)\}$i=1..P , but note that the annotation in the test set is used only for the purpose of methodology evaluation. Each annotation y represents F multi-class and binary problems, so y = $[y_1, ..., y_L] \in \{0, 1\}$N, where each problem is denoted by $y_l \in \{0, 1\}|y_l|$ | (with |yl| denoting the dimensionality of yl, where binary problems have |yl| = 1, and multi-class problems have |yl| > 1 and kylk1 = 1), and N represents the dimensionality of the annotation vector (i.e.,PL_l=1 $|y_l|$ = N). Binary problems involve an annotation that indicates the presence or absence of a class, while multi-class annotation regards problems that one (and only one) of the possible classes is present. Following the notation introduced by Estrada et al., who applied random walk algorithms for the problem of image de-noising, let us define a random walk sequence of k steps as $S_{r,k}$ = [(x(r,1), y(r,1)), ..., (x(r, k), y(r, k))], where each x(r, l) belongs to the training set D, and r indexes a specific random walk. Our goal is to estimate the probability of annotation y for a test image ex, as follows:

$$p(y|\bar{x}) = \frac{1}{Z_T} \sum_{r=1}^{R} \sum_{k=1}^{K} p(S_{r,k}|\bar{x})^{\frac{1}{k}}\, p(y|x_{(r,k)})$$
………… (1)

In (1), Z_T is a normalization factor, p(y| x(r, k)) = δ(y − y(r, k)) (with δ(.) being the Dirac delta function, which means that this term is one when y = y(r, k)), the exponent 1/ k means that steps taken at later stages of the random walk have higher weight,

$p(S_{r,k}|\bar{x})$ = p ([[$(x_{(r,1)}, y_{(r,1)})$, …., p([[$(x_{(r,k)}, y_{(r,k)})$]]| $\bar{x}$)

= $\prod_{j=2}^{k} p(x_{(r,j)}|x_{(r,j-1)}, \bar{x})\, p(y_{(r,j)}|y_{(r,j-1)})$ .. (2)

= p$(x_{r,1}|\bar{x})$p $(y_{(r,1)})$

with the derivation made assuming a Markov process and that the training labels and features are independent given the test image, p(y(r,j)|y(r,j−1)) = sy(y(r,j), y(r,j−1)) with sy(y(r,j), y(r,j−1)) = 1Zy PM m=1 m × y(r,j)(m) × y(r,j−1)(m) (_m is the weight associated with the label y(m) ∈ {0, 1} and Zy is a normalization factor), p(y(r,1)) = 1M , and p(x(r,j)|x(r,j−1), ex) and p(x(r,1)|ex) are defined.

We propose the use of class mass normalization to determine the annotation of image ex. The class mass normalization takes into consideration the probability of a class annotation and the proportion of samples annotated with that class in the training set. Specifically, we have

$\bar{y}$=

$[\sum_{i=1}^{M} y_i\,(m) * \max\,(p(y_i\,(m)|\,\bar{x}\,) - p(y(m)), 0)]$

m=1…M ………………… (3)

Where p(y(m)) = 1/M $\sum_{i=1}^{N} y_i$ (m) (m indicates the m^th dimension of label vector y). The use of class mass normalization makes the annotation process more robust to imbalances in the training set with respect to the number of training images per visual class. Notice that by in (3) represents the confidence that the image represented by ex is annotated with the labels by (m) for m = {1, ..,N}. Finally, we further process by for multi-class problems as follows:

$\forall l \in \{1, \ldots\ldots L), with\ |y_i| > 1$

$y_l^* = \{\min\,(\lfloor \bar{y}_l / \max \bar{y}_l \rfloor, 1),$ if max $(\bar{y}_l) >$ 0.5 , …………….(4)

$\{0\}^{|y_l|}$, otherwise

and for binary problems we define:

$\forall l \in \{1, \ldots\ldots L), with\ |y_i| = 1$

$y_l^* \begin{cases} 1, & \bar{y}_{l>0.5} \\ 0, & otherwise \end{cases}$ ………….(5)

As a result, the final annotation for image $\bar{x}$ is represented by y* = [$y_1^*, \ldots\ldots y_L^*$]

5.1 Proposed graph construction

The above processing motivates us to implement an advanced approach for graph structure. The fundamental ladder of the projected approach can be exposed as follows:

Input: Data matrix X∈ $R^{D*(l+u)}$, neighborhood number k.

Output: Weight matrix W = $[x_1, x_2, \ldots. x_{l+u}] \in R^{(l+u)*(l+u)}$

Algorithm:

1. Produce an error vector ρ= [$\rho_1, \rho_2, \ldots\ldots \rho_{1+u}] \in R^{1*(l+u)}$ with every element ε_j = +∞ and initialize W as a zero matrix.
2. for each sample x_j, j =1 to l+ u, do
3. Recognize the k neighborhood set as: $N_j : x_{j1} ; x_{j2} , \ldots , x_{jk}$.
4. for every sample x_{j_i}, i=1 to k, do
5. Recreate $x_{j_i} \approx \overline{x_{j_i}} = \sum_{t:t=1,t\neq i} w_{j_t}\, x_{j_t}$ according to Eq. (2).
6. if the reconstructed error $\bar{\rho}_{j_i} = \|x_{j_i} - \bar{x}_{j_i}\|_F^2 < \rho_j$, do
7. $\rho_{j_i} \leftarrow \bar{\rho}_{j_i}$, clear the j^th column of W and update it by W_{j_t}, t=1 to k , t≠ i, which are obtained in Step 5.
8. End for
9. End for
10. Output weight matrix W

Here, an effortless example of the projected label propagation procedure with outlier finding can be seen in the following Algorithm:

Input: Data matrix X∈ $R^{D*(l+u)}$, label matrix Y ∈ $R^{(1+c)*(l+u)}$, the number of nearest neighbor k and other relative parameters.

Output: The predicted label matrix F ∈ $R^{(1+c)*(l+u)}$

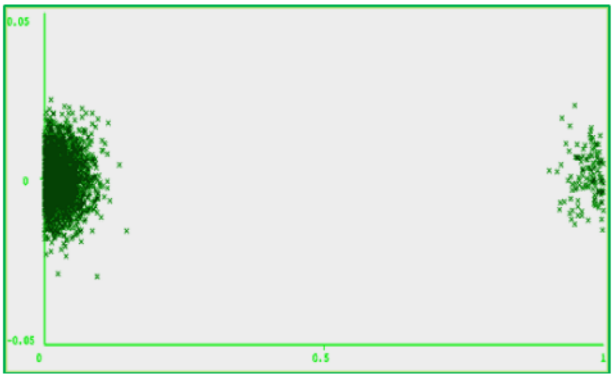
Algorithm:

1. Make the neighborhood graph and calculate the weight matrix W as Table 1.
2. Symmetries and normalize W as $\bar{W}$ = W $\Delta^{-1} W^T$ in Eq. (3), where D is the diagonal matrix fulfilling Δ_{jj} = $\sum_{i=1}^{l+u} W_{ij}$.
3. Measured the predicted label matrix F = $Y I_\beta (I - \bar{W} I_\alpha)^{-1}$ in Eq. (6) and output F.

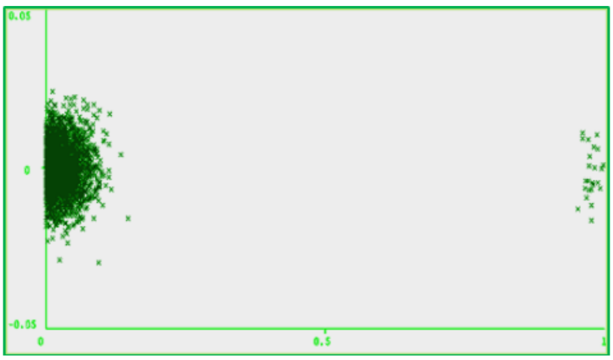
In the next section, we conducted extensive experiments on image databases.

5.2 RESULTS AND DISCUSSIONS

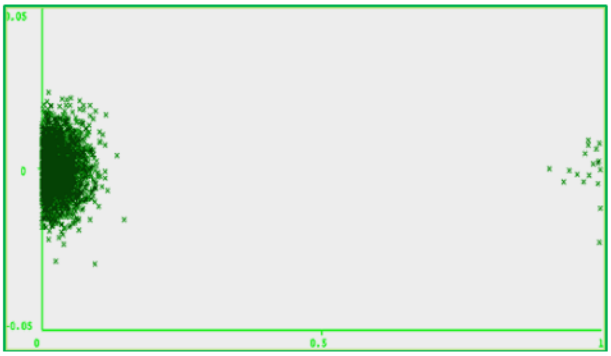
To assess the performance of the proposed graph-based web image annotation method, we conducted extensive experiments on five image databases. We then construct five image databases as follows: 1) the 6000-image database containing COREL images; 2) the 2000-image database containing Flickr images; 3) the 8000-image database containing 6000 COREL images and 2000 Flickr images; 4) the 12000-image database containing 6000 COREL, 2000 Flickr, and 4000 online images; 5) the 22000-image database containing NUSWIDE images. Every image in the database is represented by a 100-dimensional low-level visual characteristic vector and a high-level semantic feature vector, whose dimensionality is known after the training stage. Search indexes were constructing discretely over text documents from ground truth data and annotated images. The outcomes show both the effectiveness and efficiency of the proposed method and its capability to contract with the noise in the training labels. Thus the annotated text documents are able to duplicate text retrieval performance up to an accuracy of 77%. The Implementation of Graph-based annotation system using five image databases is in the following figures:



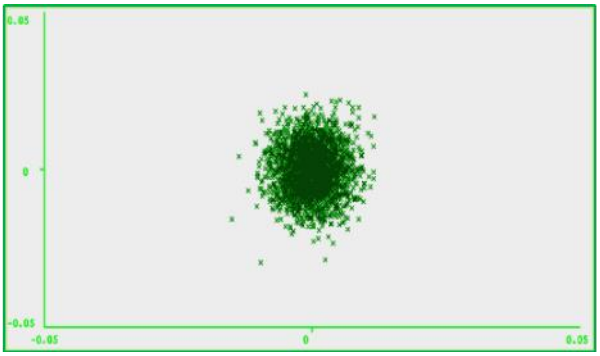
(a)



(b)



(c)



(d)

Figures- Implementation of Graph-based annotation system using five image databases.

6. CONCLUSION

In this work, we subjugated the problem of annotating images by over Corel images and their connected noisy tags. A novel graph-based Web Image Annotation for Large Scale Image Retrieval method was proposed to get better the accuracy of the image annotation. To assess the performance of the proposed method, we conducted widespread experiments on Corel image dataset. We target at solving the automatic image annotation issue in a new search and mining framework. First in searching stage, it detects a set of visually alike images from a large-scale image database which are considered useful for labeling and retrieval. Then in the mining stage, a graph pattern matching algorithm is used to discover most representative keywords from the annotations of the retrieved image subset. To be able to rank relevant images, this approach is extensive to Probabilistic Reverse Annotation. The outcomes demonstrate both the effectiveness and efficiency of the proposed approach and its ability to deal with the noise in the training labels.

REFERENCES

[1] Yu Tang Guo and Bin Luo," An automatic image annotation method based on the mutual K-nearest neighbor graph", IEEE Natural Computation (ICNC), 2010 Sixth International Conference on 10-12 Aug. 2010, DOI: 10.1109/ICNC.2010.5584164.

[2] Xirong Li, Le Chen and Lei Zhang," Image Annotation by Large-Scale Content-based Image Retrieval", MM'06, October 22–27, 2006, Santa Barbara, CA, USA. Copyright 2006 ACM 1-58113-000-0/00/0004.

[3] C. Lakshmi Devasena and T. Sumathi," An Experiential Survey on Image Mining Tools,Techniques and Applications", International Journal on Computer Science and Engineering (IJCSE), ISSN : 0975-3397 Vol. 3 No. 3 Mar 2011.

[4] Pramod Sankar K. and C. V. Jawahar," Probabilistic Reverse Annotation for Large Scale Image Retrieval", IEEE Conference on Computer Vision and Pattern Recognition, 2007. 17-22 June 2007, DOI: 10.1109/CVPR.2007.383169.

[5] A. Balasubramanian, Million Meshesha, and C. V. Jawahar. Retrieval from document image collections. In 7th International Workshop on Document Analysis Systems, DAS, pages 1.12. LNCS, Springer-Verlag, 2006.

[6] Rajashree Galgali and Rahul Joshi," A Survey: Image Indexing and Retrieval Using Automated Annotation", International Journal For Innovative Research In Multidisciplinary Field Issn – 2455-0620 Volume - 3, Issue – 6, June – 2017.

[7] Mingbo Zhao and Tommy W.S. Chow," Automatic image annotation via compact graph based semi-supervised learning", Knowledge-Based Systems 76 (2015) 148–165.

[8] Keiji Yanai and Kobus Barnard," Finding Visual Concepts by Web Image Mining", ACM 1595933329/ 06/0005, May 23–26, 2006.

[9] Barbora Zahradnikova and Sona Duchovicova," Image Mining: Review and New Challenges", (IJACSA) International Journal of Advanced Computer Science and Applications,Vol. 6, No. 7, 2015.

[10] Xiaojun Qi, and Ran Chang," A Scalable Graph-Based Semi- Supervised Ranking System for Content-Based Image Retrieval", International Journal of Multimedia Data Engineering and Management, 4(4), 15-34, October-December 2013.

[11] R. Datta, D. Joshi, J. Li, and J. Wang. Image retrieval: Ideas, influences, and trends of the new age. ACM Comput. Surv., 40(2), 2008.

[12] J. Wang, F. Wang, C. Zhang, H.C. Shen, L. Quan, Linear neighborhood propagation and its applications, IEEE Trans. Pattern Anal. Mach. Intell. 31(9) (2009) 1600–1615.

[13] X. Zhu, Z. Ghahramani, J.D. Lafferty, Semi-supervised learning using Gaussian fields and harmonic function, in: Proc. of ICML, 2003.

[14] Liana Stanescu and Dumitru Dan Burdescu," Medical Images Annotation", Springer, New York, NY, 2011. https://doi.org/10.1007/978-1-4614-1909-9_5.

[15] Wang Y (2008) Automatic image annotation and categorization, PhD thesis, University of London, September.

[16] Jing-Ming Guo, Senior Member, IEEE, and Heri Prasetyo, "Content-Based Image Retrieval Using Features Extracted From Halftoning-Based Block Truncation Coding", IEEE Transactions on Image Processing, VOL. 24, NO. 3, MARCH 2015.

[17] Baoyuan Wu a, SiweiLyu b, Bao-GangHu a, QiangJi,"Multi-label learning with missing labels for image annotation and facial action unit recognition",2015ElsevierLtd.

[18] Jing-Ming Guo, Senior Member, IEEE, Heri Prasetyo, and Jen-Ho Chen,"Content-Based Image Retrieval Using Error Diffusion Block Truncation Coding Features",IEEE Transactions On Circuits And Systems For Video Technology, VOL. 25, NO. 3, MARCH 2015.

[19] Marina Ivasic-Kos a,n, MiranPobar a, SlobodanRibaric b,"Two-tier image annotation model based on a multilabel classifier and fuzzy-knowledge representation scheme",2015ElsevierLtd.

[20] Bin Xu, Student Member, IEEE, Jiajun Bu, Member, IEEE, Chun Chen, Member, EEE,"EMR: A Scalable Graph-based Ranking Model for Content-based Image Retrieval",IEEE Transactions on Knowledge and Data Engineering, VOL. 27, NO. 1, JANUARY 2015.