ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

An Effective Survey on Mining Weakly Labelled Web Facial Images for Search-Based Face Annotation

Pooja P. Gadhave

Final Year M.E. Student, Department of Computer, Dattakala Group of Institutions college of Engineering, Swami-Chincholi, Taluk-Daund, Pune, India- 413801

ABSTRACT: Developments in the area of image mining have shown the way for incredible growth in extensively large and detailed image databases. The images which are available in databases, if checked, can endow with valuable information. The paper investigates framework of search-based face annotation (SBFA) by mining weakly labeled facial images that are freely available on the World Wide Web (WWW). One challenging problem for the search-based face annotation scheme is how to effectively perform annotation by exploiting list of most similar facial images and their weak labels that are often noisy and incomplete. To tackle this problem, we propose effective unsupervised label refinement (ULR) approach for the refining the labels of web facial images using machine learning techniques. We formulate a learning problem as a convex optimization and develop effective optimization algorithms to solve the large-scale learning task efficiently. To further speed up the proposed scheme, we also propose a clustering-based approximation algorithm which can improve scalability considerably. We have conducted an extensive set of empirical studies on a large-scale web facial image tested, in which encouraging results showed that the proposed ULR algorithms can significantly boost the performance of the promising SBFA scheme

KEYWORDS: Face annotation, web facial images, search base face annotation, content-based image retrieval, weak Label, Search based facial annotation.

I. INTRODUCTION

Digital media devices are increasing so the different social media tools used for sharing photos. The large number of human facial images shared over the different social real world application some of this images are tagged properly but many of images are not tagged properly so the facial annotation are came.

Auto face annotation can be beneficial to the many real world applications. For example, the auto face annotation techniques and online photo-sharing sites (e.g., Face book) can automatically annotate users' uploaded photos to the facilitate online photo search and management. Besides, face annotation can also be applied in news video domain to detect important persons appeared in the videos to help news video retrieval and summarization tasks. Classical face annotation approaches are often treated as an extended face recognition problem, where different classification models are trained from a collection of well labeled facial images by employing the supervised/semi-supervised machine learning techniques. However, the "model-based face annotation" techniques are limited in the several aspects. First, it is usually time-consuming and expensive to collect large amount of human-labeled training facial images. Second, it is usually difficult to the generalize the models when new training data or new persons are added, in which an intensive retraining process is usually required. Last but not least, the annotation/recognition performance often scales poorly when the number of persons or classes is very large. Recently, some emerging studies have attempted to the explore a promising search-based annotation paradigm for facial image annotation by mining the World Wide Web (WWW), where a massive number of weakly labeled facial images are freely available. Instead of the training explicit classification models by the regular model-based face annotation approaches, the search-based face annotation (SBFA) paradigm aims to tackle the automated face annotation task by exploiting content-based image retrieval (CBIR) techniques in mining massive weakly labelled facial images on the web. The SBFA framework is data-driven and model-free, which to some extent is inspired by the search-based image annotation techniques for generic image annotations. The main objective of the SBFA is to assign correct name labels to a given query facial image. In particular given a novel facial image for annotation, we first retrieve a short list of top K most similar facial images

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

from a weakly labeled facial image database, and then annotate the facial image by performing voting on the labels associated with the top K similar facial images.

One challenge faced by such SBFA paradigm is how to effectively exploit short list of candidate facial images and their weak labels for the face name annotation task. To tackle above problem, we investigate and develop a search-based face annotation scheme. In particular, we propose a novel unsupervised label refinement (URL) scheme by exploring machine learning techniques to enhance the labels purely from weakly labelled data without human manual efforts. We also propose the clustering-based approximation (CBA) algorithm to improve the efficiency and scalability. As a summary, the main contributions of the paper include the following:

.

- We investigate and implement a promising search based face annotation scheme by mining the large amount of weakly labeled facial images freely available on the WWW.
- We propose a novel ULR scheme for enhancing the label quality via a graph-based and low-rank learning approach.
- We propose an efficient clustering-based approximation algorithm for the large-scale label refinement problem.
- We conducted an extensive set of experiments, in
  Which encouraging the results were obtained.

We note that short version of this work had appeared in SIGIR2011. This journal article has been significantly extended by including the substantial amount of new content.

The remainder of this paper is organized as follows: Section 2 reviews of the related work. Section 3 gives an overview of proposed search-based face annotation framework. Section 4 presents proposed unsupervised label refinement scheme. Section 5 shows our experimental results of the performance evaluation, and Section 6 discusses the limitation of our work. Finally, Section 7 concludes of this paper.

II. RELATED WORK

Different studies are perform on face annotation in mining weakly labeled facial images which are present over internet in this human name are treated as input query and aim is to refine the text-based search results by achieving consist facial images. Group of related work is on the topics of face recognition and verification, which are classical research problems in the computer vision and pattern recognition and have been extensively studied for many years. Recent years have observed some emerging benchmark studies of unconstrained face detection and verification techniques on the facial images that are collected from the web, such as the LFW benchmark studies. Some recent study had also attempted to extend classical the face recognition techniques for face annotation tasks. Comprehensive reviews on face recognition and verification topics can be found in the some survey papers.

The second group is about studies of generic image annotation. The classical image annotation approaches as usually apply some existing object recognition techniques to train classification models from human-labelled training images or attempt to infer the correlation/probabilities between images and annotated keywords. Given limited training data and semi-supervised learning methods have also been used for image annotation. For example, Wang et al. Proposed to refine the model-based annotation results with label similarity graph by following random walk principle. Similarly, Pham et al. proposed to annotate unlabeled facial images in the video frames with an iterative label propagation scheme. Although semi-supervised learning approaches could leverage both labelled and unlabeled data and it remains fairly time-consuming and expensive to collect enough well-labeled training data to achieve good performance in large-scale scenarios. Recently search-based image annotation paradigm has attracted more and more attention. For example, Russell et al. Built a large collection of web facial images with ground truth labels to facilitate object recognition research. However, most of the works were focused on the indexing, search, and feature extraction techniques. Unlike these existing works, we propose the novel unsupervised label refinement scheme that is focused on optimizing the label quality of facial images towards the search-based face annotation task.

The third group is about face annotation on personal or family or social photos. Several studies have mainly focused on the annotation task on personal photos, which often contain rich contextual clues, such as personal/family names, social context, geotaxis, time stamps and so on. The number of persons or classes is usually quite small, making such annotation tasks less challenging. These techniques usually achieve fairly accurate annotation results, in which some techniques have been successfully deployed in commercial applications, for example Apple, iPhoto, Google Picasa and Microsoft easy Album and Face book face auto tagging solution.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

The fourth group is about studies of face annotation in mining weakly labelled facial images on the web. Some studies consider a human name as a input query, and mainly aim to refine the text-based search results by exploiting visual consistency of facial images. For example Ozkan and Duygulu proposed a graph-based model for finding a densest sub-graph as the most related result.

For example, Berg et al proposed a possibility model combined with a clustering algorithm to estimate the relationship between the facial images and the names in their captions. For the facial images and the detected names in the same document (a web image and its corresponding caption), Guillaumin etal proposed to iteratively update the assignment based on a minimum cost matching algorithm. In their follow-up work, they further improve the annotation performance by using distance metric learning techniques to achieve more discriminative feature in low-dimension space.

Our work is different from the above previous works in two main aspects, First our work aims to solve the general content-based face annotation problem using the search-based paradigm, where facial images are directly used as query images and the task is to return the corresponding names of the query images. Very limited research progress has been reported on this topic. Some recent work mainly addressed the face retrieval problem, in which an effective image representation has been

proposed using both local and global features. Second, based on the initial weak labels, the proposed unsupervised label refinement algorithm learns an enhanced new label matrix for all the facial images in the whole name space; however, the caption-based annotation scheme only considers the assignment between the facial images and the names appeared in their corresponding surrounding-text.

As a result, the caption-based annotation scheme is only applicable to the scenario where both images and their captions are available, and cannot be applied to the our SBFA framework due to lack of complete caption information.

The fifth group is about the studies of the purifying web facial images, which aims to leverage noisy web facial images for face recognition applications. Usually these works are proposed as a simple pre-processing step in whole system without adopting sophisticated techniques.

For example, the work in applied a modified means clustering approach for cleaning up the noisy web facial images. Zhao et al. proposed a consistency learning method to train face models for the celebrity by mining the text-image co occurrence on the web as a weak signal of relevance toward supervised face learning task from a large and noisy training set. Unlike the above existing works, we employ the unsupervised machine learning techniques and propose a graph-based label refinement algorithm to optimize the label quality over a whole retrieval database in SBFA task.

Finally, we note that our work is also related to our recent work of the WLRLCC method in and our latest work on the unified learning scheme. Instead of enhancing the label matrix over the entire facial image database, the WLRLCC algorithm is focused on learning more discriminative features for the top retrieved facial images for each individual query, which thus is very different from the ULR task in this paper. Last but not least, we note that the learning methodology for the solving unsupervised label refinement task are partially inspired by some existing studies in machine learning, including graph based semi-supervised learning and multilabel learning techniques.

SEARCH-BASED FACE ANNOTATION

System Architecture:

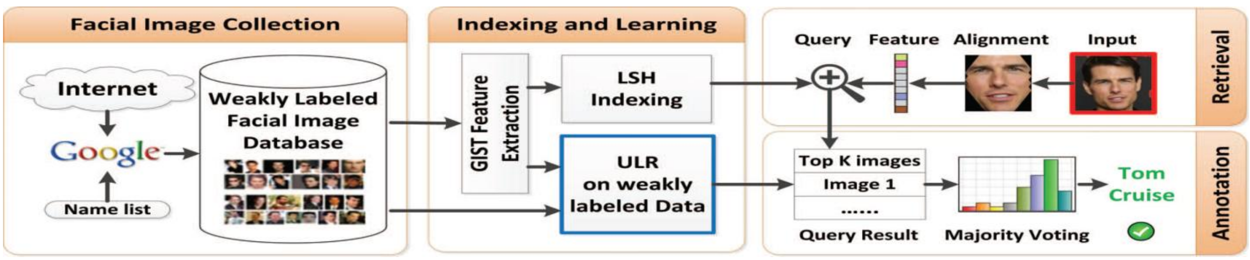


Fig. 1. The system flow of the proposed search-based face annotation scheme. (a) We collect weakly labeled facial images from WWW using web search engines. (b) We preprocess the crawled web facial images, including face detection, face alignment, and feature extraction for the detected faces; after that, we apply LSH to index the extracted high-dimensional facial features. We apply the proposed ULR method to refine the raw weak labels together with the proposed clustering-based approximation algorithms for improving the scalability. (c) We search for the query facial image to retrieve the top K similar images and use their associated names for voting toward auto annotation.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 5, Issue 5, May 2017

In this section, it briefly introduce the framework of SBFA system. Figure 1 illustrates the proposed framework that consists of the following four major modules:

(1) The database construction module by crawling facial images from the World Wide Web;
(2) The database indexing module towards fast retrieval of high-dimensional facial features;
(3) The content-based facial image retrieval module for Searching a query facial image;
(4) The automated face annotation module for naming the query by mining the top-k retrieved similar facial images and their corresponding weakly labels.

The first four steps are usually conducted before the test phase of the face annotation task, while the last two steps are conducted during the test phase of a face annotation task, which usually should be done very efficiently. It can annotate the query facial image by exploiting both weakly label information and visual contents of top-ranked facial images to maximize the annotation efficiency.

Face Recognition: The face recognition problem can be divided into the two main stages: The face verification (authentication) and face identification (recognition). The detection stage is the first stage; it includes identifying and locating the face in an image. The recognition stage is second stage, it includes feature extraction, where important information for discrimination is saved, and the matching, where the recognition result is given with the aid of a face database. The several face recognition methods have been proposed. In the literature on the topic there are different classifications of the existing techniques. The following is one of the possible high-level classification:

Holistic Methods- The whole face image is used as the raw input to recognition system. An best example is the well-known PCA-based technique.

Firstly local features are extracted, such as eyes, nose and mouth. Their locations and the local statistics (appearance) are the input to the recognition stage. Pre-processing: It pre-process facial images to extract face related information, including face detection and alignment, facial region extraction, and facial feature representation. As a result, each face can be represented by a d-dimensional feature vector. An object recognizer using PCA (Principle Components Analysis):-

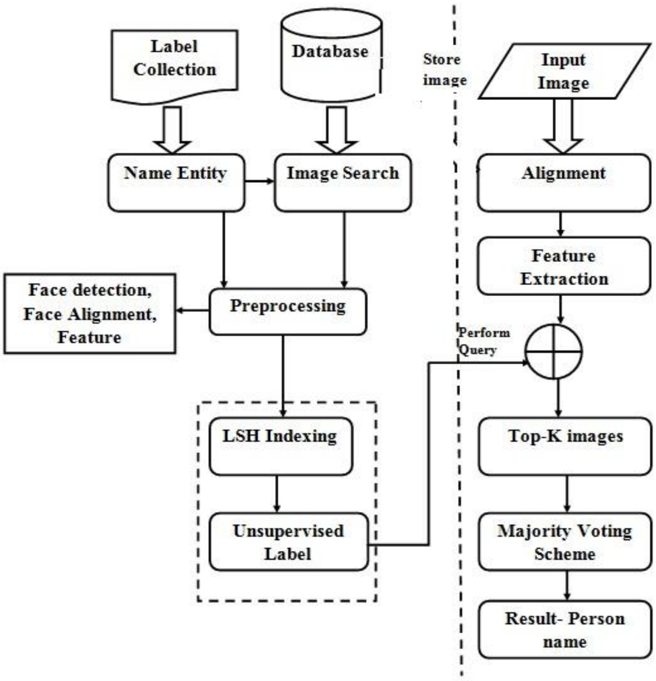


Figure2: Search Based Face Annotation System Flow

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 5, Issue 5, May 2017

III. PROPOSED ALGORITHM

System Algorithms:

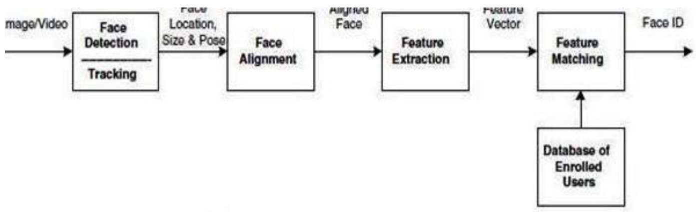


Figure 3: Face recognition processing flow

1. Get the Eigen vectors that form the Eigen space
2. Get or set the labels for the corresponding training image
3. Get or set the Eigen distance threshold.
4. Get the average Image.
5. Get the Eigen values of each of the training image
6. Create an object recognizer using the specific training data and parameters, it will always return the most similar object.
7. Create an object recognizer using the specific training data and parameters
8. Given the Eigen value, reconstruct the projected image.
9. Get the Euclidean Eigen-distance between input image and every other image in the database
10. Given the image to be examined, find in the database the most similar object, return the index and the Eigen distance
11. Try to recognize the image and return its label.

IV. PSEUDO CODE

Unsupervised Label Refinement:

It is key step of the framework is to engage an unsupervised learning scheme to enhance the label quality of the weakly labelled facial images. This process is very important to the entire search based annotation framework since the label quality plays a critical factor in final annotation performance. It denoted by X IRnxd the extracted facial image features, where n and d represent the number of facial images and the number of feature dimensions, respectively. Further it denoted by Ω= {n1, n2, . . .,; nm} the list of human names for annotation, where m is the total number of human names. It denotes Y [0; 1] nxm as the initial raw label matrix to describe the weak label information, in which the ith row Y represents the label vector of the ith facial image X IRd. In this application, Y is often noisy and incomplete.

In particular, for each weak label value Yij, Yij ≠ 0

indicates that the ith facial image Xi has the label name n,while Yij ≠ 0 indicates that the relationship between ith facial image Xi and jth name is unknown. Note that it usually have Yi* = 1 since each facial image in database was uniquely collected by a single query. Following the terminology of graph-based learning methodology, build a sparse graph by computing the weight matrix W = [Wij] IRnxn, where Wij represents the similarity between Xi and Xj.

Clustering Base Approximation:

To further enhance the Scalability and efficiency in algorithms, in this paper, proposed a clustering-based approximation solution to speed up the solutions for large-scale problems. In the particular, the clustering strategy could be applied in two different levels: 1) one is on “image-level,” which can be used to directly separate all the n

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

facial images into a set of clusters, and 2) the other is on "name-level," which can be used to First separate the m names into a set of clusters, then to further split the retrieval database into different subsets according to the name-label clusters. Typically, the number of facial images n is much larger than the number of names m, which means that the clustering on "image level" would be much more time-consuming than that on "name-level." Thus, this approach, adopt the "name level" clustering scheme for the sake of the scalability and the efficiency. After the clustering step, solve the proposed ULR problem in each subset, and then merge all the learning results into the final enhanced label matrix F.

Algorithm:
Input: c, qc, Lloop
Output: clustering highest order result list Llist
Add M0 to Llist;
Repeat
Remove the largest cluster Ml from Llist;
For i=1 to t do
Bisect Ml to M1(i) and M2(i);
Compute sum of squared error (SSEi);
Select the result with the lowest SSEi value;
Add m1(i),m2(i) to Llist;
Until ¦Llist¦=qc;

In the Clustering scheme, the ith row Ci* is used as the feature vector for class Xi. In each step, the largest cluster is bisected for I loop times and the clustering result with the lowest sum-of-square-error (SSE) value is used to update the clustering lists. In our framework, we set Iloop to 10. The details of the Bisect Clustering Based Approximation (BCBA) scheme are illustrated in above algorithm, where qc is the cluster number.

Weighted Majority Voting (WMV): It is used to combines the set of labels associated with these top K similar face examples. It is based on top-n retrieval images. The confidence weight depends on the Euclidean distance between the query image and the similarity image. In particular, for the i-th nearest similar face, it assigns a weight coefficient wi to the corresponding label vector yi by wi =

$$\frac{\phi(X_q, X_i)}{\sum_{j=1}^{n} \phi(X_q, X_j)}$$

where (·, ·) is related to the distance between the query image xq and its i-th nearest sample xi .

$$\phi(X_q, X_i) = \frac{1}{1+\exp(-\|X_q - X_i\|^2)}$$

V. SIMULATION RESULTS

Evaluation of Label Enhancement :

This experiment aims to evaluate the performance of the refined label matrix Y. It presents the results of the proposed SBFA algorithm on the most noisy DB-100K database. Similar observations can also been observed on the other databases by employing different algorithms. The ULR, CBA algorithm is applied to learn and enhance the label matrix. After that, both the refined label matrix Y and the initial weak label matrix are used in the face name annotation step. The comparison results are presented in Figure 3, it represents evaluation of the enhanced label matrix Y on the DB-100K database.

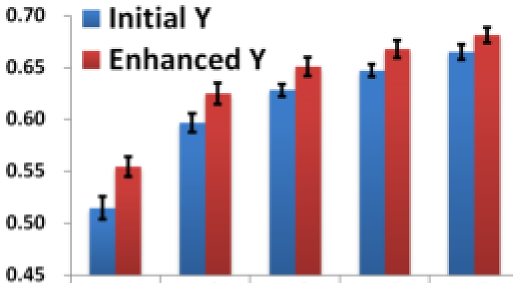
ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017



	t=1	t=2	t=3	t=4	t=5
Initial Y	0.5145	0.5961	0.6283	0.6471	0.6649
Enhanced Y	0.5544	0.6248	0.6509	0.6679	0.6813

Figure 3: Evaluation of the enhanced label matrix Y on the DB-100K database.

Next experiment is to demonstrate the qualitative tagging performance achieved by different SBFA and ULR methods for automatic image tagging tasks. To achieve this purpose, it randomly select several images from the test set, and applied a number of different ULR methods to annotate them using the proposed search based annotation approach Evaluation of Varied k Values:

Figure 4 shows the performance of SBFA at top t tags by varying k, the number of top retrieved similar images from 10 to 60. From the results, observed that k affects the annotation performance. In particular, when k is about 40 to 50, the proposed method achieved the best average precision. This is reasonable because if k is too small, some relevant images may not be retrieved, while if k is too large, lots of irrelevant images could be retrieved, leading to engage many noisy tags in the list of candidate tags. Both of the above situations could degrade the annotation performance.

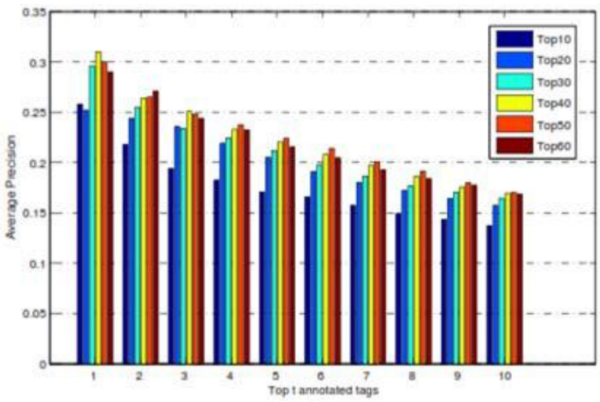


Figure 4: Comparisons of average precision under different top k similar images used.

Result Images

Figure 5: Image Upload

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

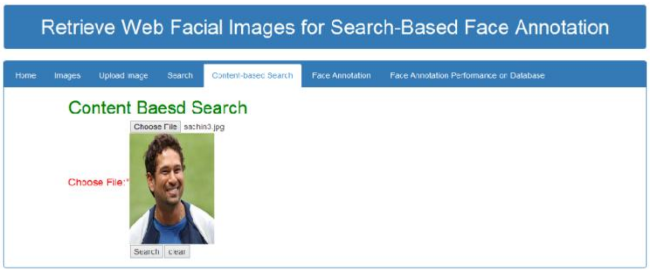


Figure 6: Content Based Search

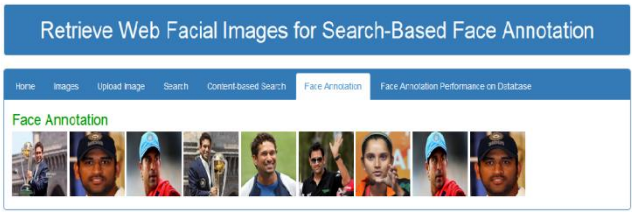


Figure 7: Face Annotation

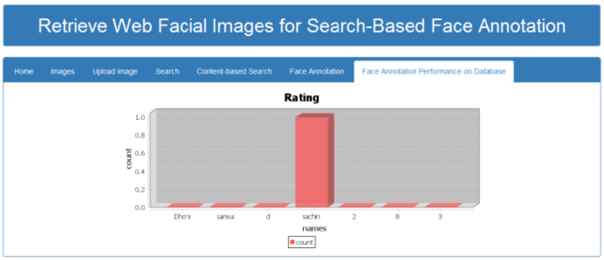


Figure 8: Face Annotation Performance on Database

Limitations

Work is limited in several aspects. First, we assume each name corresponds to a unique single person. Duplicate name can be a practical issue in real-life scenarios. One future direction is to extend our method to address this practical problem. Second, we assume the top retrieved web facial images are related to a query human name. This is clearly true for celebrities. However, when the query facial image is not a well-known person, there may not exist many relevant facial images on the WWW, which thus could affect the performance of the proposed annotation solution. This is a common limitation of all existing data-driven annotation techniques.

VI. CONCLUSION

It investigate a promising search-based face annotation framework, in which it focused on tackling the critical problem of enhancing the label quality and proposed a ULR algorithm. To further improve the scalability, it also proposed a clustering-based approximation solution, which successfully accelerated the optimization task without introducing much performance degradation. It uses unsupervised labeled refinement (ULR) approach for improving efficiency and scalability of proposed system. Results obtained shows that the proposed ULR technique can significantly boost the performance of the promising search-based face annotation scheme.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

REFERENCES

[1]. Y. Tian, W. Liu, R. Xiao, F. Wen, and X. Tang, "A Face Annotation Framework with Partial Clustering and Interactive Labeling," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2007.

[2]. Dayong Wang, Steven C.H. Hoi, Member, IEEE, Ying He, and Jianke Z "Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation" IEEE transactions on knowledge and data engineering, vol. 26, no. 1, January 2014.

[3]. A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-Based Image Retrieval at the End of the Early Years," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.

[4]. S.C.H. Hoi, R. Jin, J. Zhu, and M.R. Lyu, "Semi-Supervised SVM Batch Mode Active Learning with Applications to Image Retrieval," ACM Trans. Information Systems, vol. 27, pp. 1-29, 2009.

[5]. X.-J. Wang, L. Zhang, F. Jing, and W.-Y. Ma, "AnnoSearch: Image Auto-Annotation by Search," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR), pp. 1483- 1490, 2006.

[6]. W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," ACM Computing Survey, vol. 35, pp. 399- 458, 2003.

[7]. Z. Cao, Q. Yin, X. Tang, and J. Sun, "Face Recognition with Learning-Based Descriptor," IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 2707-2714, 2010.

[8]. R. Jafri and H.R. Arabnia, "A Survey of Face Recognition Techniques," J. Information Processing Systems, vol. 5, pp. 41-68, 2009.

[9]. D. Anguelov, K. Chih Lee, S.B. Go¨ktu¨ rk, and B. Sumengen, "Contextual Identity Recognition in Personal Photo Albums," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR '07),2007.

[10] L. Wu, S.C.H. Hoi, R. Jin, J. Zhu, and N. Yu, "Distance Metric Learning from Uncertain Side Information for Automated Photo Tagging," ACM Trans. Intelligent Systems and Technology, vol. 2, no. 2, p. 13, 2011.

[11] P. Wu, S.C.H. Hoi, P. Zhao, and Y. He, "Mining Social Images with Distance Metric Learning for Automated Image Tagging," Proc. Fourth ACM Int'l Conf. Web Search and Data Mining (WSDM '11), pp. 197-206, 2011.

[12] D. Wang, S.C.H. Hoi, and Y. He, "Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation," Proc. 34th Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2011.

BIOGRAPHY

Pooja Prakash Gadhave is student and appeared for final year, Master of Engineering (M.E) in Department of Computer, Dattakala Group of Institutions college of Engineering, Swami-Chincholi,Taluk-Daund, Pune, She received Master of Computer Science and Engineering (ME) degree in 2017 from Pune University, India. Her research interests are Data Mining, Unsupervised Label Refinement Algorithms, Face Recognition, Clustering Base Approximation etc.