Volume 4, Issue 11, November 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Survey on Search Based Face Annotation Using Weakly Labelled Facial Images

Shital A. Shinde*, Prof. Archana Chaugule
Department of Computer Engg, DYPIET Pimpri,
Savitri Bai Phule Pune University, Maharashtra India

Abstract— Face recognition/detection presents a challenging problem in the field of image analysis and computer vision, and it is becoming more popular day by day because of its applications in various domain areas. Face annotation is one of the ways for recognizing face images. Auto face annotation is playing important role in many real-world applications. Face annotation task is part of face detection and recognition. Now a day's research interest in mining weakly-labelled facial images on the internet for resolving research challenge in computer vision and image understanding is growing. A content-based image retrieval (CBIR) system needs users to query images by their low quality visual content. This process not only makes it difficult for users to formulate queries, but also can lead to unsatisfied results of image retrieval. To tackle this problem, Image annotation technique is used. The main aim of image annotation process is to automatically assign associate label to images, so image retrieving users are able to query images by labels and automatically detect human faces from a photo image and further name the faces with the corresponding human names. This paper provides various techniques or methods that are used for mining facial images from internet and annotating facial images.

Keywords— Face annotation, Search based face annotation, unsupervised label refinement, web facial images, and weak label.

I. INTRODUCTION

Day by the digital accessories for capturing images is increasing and sources for sharing that images e.g., social media tools/sites are also increasing rapidly. Large number of images is shared over these social sites, but many times images shared by person are not having any label so it becomes problematic in understanding name of person from image if any random person sees it. Tagging facial images are known as face annotation and now a day many techniques are introducing for annotation. Auto face annotation is used for automatic face image annotation without any human intervention [1]. Facial annotation is also applying for videos, such as annotation of facial images from news video is done and then it showed on television so that peoples can recognize person in TV [2].

The model base annotation has more limitations i.e. it is more time consuming and more costly to collect large amount of human labelled training facial image. It is more difficult to generalize the models when new persons are added, in which retraining process is required and last the annotation performance is become poor when the number of persons is very more. To address the challenges "Auto face annotation" is important technique which automatically gives name to relevant person images. This technique is more beneficial to different real world application for (e.g. Facebook) which annotates photos uploaded by the users for managing online album and searches the photos. Search base annotation is used for facial image annotation by mining the World Wide Web (WWW), where large numbers of weakly labelled facial Images are freely available. The automated face annotation task is accepted as challenging aim of the search based face annotation by using content-based image retrieval (CBIR) techniques in mining number of weakly labelled facial images on the web [3, 4].

The main objective of search-based face annotation is to assign correct name labels to a given query facial image [16]. The Search Based Face Annotation (SBFA) framework is data driven as well as model free, which is inspired by search based image annotation methods of image annotation. The main aim of SBFA is assign correct associative name label to image provided as input to search query. In this approach top k most similar facial images are retrieved from database, and respectively task of annotation is performed on the basis of majority voting scheme. For effective name labelling annotation task, a novel unsupervised label refinement (ULR) scheme is used which solves problem of weakly label face annotation. Clustering-based approximation (CBA) algorithms are also exploited in this work, which lead to improvement of efficiency and scalability of search based system [16].

Study of this work is related to several groups of research work, where work is mostly carried for facial images recognition, feature exaction, annotation etc. Various groups studied in these work are face recognition and verification, generic image Annotation, face annotation on personal/ family/social photos, face annotation in mining weakly labelled facial images on the web, and purify web facial images. This paper, aims study of various techniques used for face annotation.

II. LITERATURE SURVEY

Various techniques are present for face annotation in mining weakly labelled facial images from internet. As per the study it shows that most of the techniques accepts name of person as input and process text-based search for achieving facial image. Many researchers are trying to propose system, which will accept image-based input and generate text-based output. Various research groups are working for successfully fulfil this objective.

A. Face recognition and verification

These are classical research problems in computer vision and pattern recognition and have been studied from many years. G.B. Huang et al. [5] designed Labelled Faces in the Wild (LFW), is designed to address the problems, although it can be used to address the others if desired. Proposed work referred to this problem as the pair matching problem. This work describes the details of the database and its acquisition; it provides specific experimental paradigms for suitable database. Their work is carried with consistent and comparable database for performing research. They have introduced a new database, Labelled Faces in the Wild, whose primary aim is to provide a large database of real world face images for the unseen pair matching problem of face recognition, and fit it neatly into the detection-alignment-recognition pipeline.

Z. Cao et al. [6] presented a novel approach to address the representation and the matching issue in face recognition. In proposed work they firstly worked on approach that encodes the micro-structures of the face by a new learning-based encoding method. They used unsupervised learning techniques to learn an encoder from the training sets. In next step they applied PCA to get a compact face descriptor. From generated result, proposed work suggested that the discriminative ability of the descriptor can be improved by a simple normalization mechanism after PCA. To handle the large pose variation in real-life scenarios, they proposed a pose-adaptive matching method that uses pose-specific classifiers to deal with various pose combinations of the matching face pair. This proposed approach is comparable with the state-of-the-art methods on the Labelled Face in Wild (LFW) benchmark. Face micro-pattern encoding is learned in this work but the pattern sampling is still manually designed. Result of these work produced compact, highly discriminative, and easy-to-extract learning-based (LE) descriptor.

B. Image annotation

J. Fan, Y. Gao, and H. Luo [7] proposed a novel algorithm for automatic multi-level image annotation using hierarchical classification. A novel multi-modal algorithm is proposed in this work to achieve reliable interpretation of the contextual relationships between the atomic image concepts and the co-appearances of salient objects. The major problem with hierarchical approach is that the classification errors may be transmitted among different concept levels (i.e., *inter-level error transmission*).To tackle the inter level error transmission problem, a novel hierarchical boosting algorithm is proposed and multi-task learning is to boost hierarchical image classifier training set. Hierarchical image classifier training algorithm proposed in this paper is able to simultaneously learn the classifiers for large amounts of image concepts with concept of visual diversities and similarities. A novel hyperbolic visualization framework is seamlessly incorporated to enable intuitive query specification and similarity-based evaluation of large amounts of returned images.

Hierarchical image classification has two advantages, first is that the classifiers for the high-level image concepts can effectively be learned by combining the classifiers for the relevant image concepts at the lower levels of the concept ontology; and second is that the computational complexity for training the classifiers for large amounts of image concepts can significantly be reduced through exploiting the strong correlations between the image concepts.

X.-J. Wang et al. [8] presented AnnoSearch, a novel scheme to annotate images using search and data mining technologies. In proposed work they solved this problem in two-stages; first is searching for semantically and visually similar images on the web sites, and mining annotations from them. At least one accurate keyword is necessary for enabling text-based search for a set of semantically similar images. After that content-based search task is performed on this retrieved set to obtain visually similar images. As last processing task of their work, annotations are mined from the descriptions of these images. These proposed works has advantage that is no supervised training process is adopted in this approach, and as a result, it handles unlimited vocabulary. It also ensures a highly scalable image database. In their work they mentioned that in follow-up work they will work on reinforcing the labels of images from large scale database and they are interested to resolve the problem of how to annotate query images without associated keywords.

J. Tang et al. [9] proposed a novel kNN-sparse graph-based semi-supervised learning method with regularization on training labels, which is used to annotate noisily-tagged web images by label propagation. Here the graph is constructed to handle the semantically unrelated links. It is constructed by reconstructing each sample from its k nearest neighbours to improve the efficiency, while the approximate method is applied to accelerate the kNN search. And the regularization is proposed to handle the noise in the training labels. Experimental results showed a key factor, which affects the performance of image annotation process with the tags as training labels. Actually, for image annotation, not need to correct all the noisy tags, instead they have collected the correct image-label pairs as much as possible for training. They also decided to focus on how to construct an effective training set from the community-contributed images and tags in future work.

C. Face annotation on personal/family/social photos

Jae young choi et al. [10] they proposed a novel collaborative framework of face recognition for improving the accuracy of face annotation. Multiple FR engines available in online social networks (OSN's) are used for effective FR. This paper includes two main tasks, first is the selection of expert FR engines to recognize query face images. And

second is the merging of multiple FR results, generated from different FR engines, into single FR results. These works implement the viola-Jones face detection algorithm for detecting facial images in personal photos. But in practical perspective it becomes problematic depending on targeted application and associated parameter setup. To tackle this problem more advanced face detection technique can be used in face annotation framework, which provides more accurate results.

D. Face annotation in mining weakly labelled facial images on the web

M. Guillaumin et al. [11] introduced a modification to incorporate the constraint that a face is only depicted once in an image. This work has two scenarios of naming persons in database for finding face of person and assigning name to all faces. The text based result is not more accurate. Graph based approach is improved by introducing the constraint, objective function generative models have previously been proposed to solve the multi-person naming task by comparing generative and graph based methods. The most significant graph based method is extended in future to multi-person naming. M. Guillaumin et al. [11] proposed a method to iteratively update the assignment based on a minimum cost matching algorithm. In their follow-up work Guillaumin et al. [12], they further uses distance metric learning technique to enhance the annotation performance to gain lots of distinguish features in low dimension space.

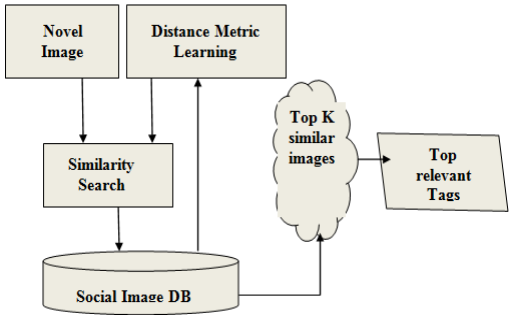

Fig 1: Retrieval-Based Annotation Approach

Fig 1. Shows that retrieval based approach are applied with distance metric learning also various different techniques are implemented with these retrieval based or search based face annotation. As shown in the Fig-2 a similarity search task is conducted, which is used to find a set of top-k images from a social or World Wide Web image database, and it obtains relevant images.

Z. Wu et al. [13] mainly addressed the face retrieval problem, by using local and global features which propose an effective image representation. Future Work visual word vocabulary for face is improved by designing a supervised learning algorithm. This system is highly scalable, and they plan by using a computer cluster to apply on a web-scale image database.

D. Wang, Steven C.H. Hoi et al. [16] proposed an effective unsupervised label refinement algorithm for refining the web facial images. Performance of the proposed system is improved using optimization algorithm which is helpful to solve large-scale learning effectively i.e. clustering based approximation algorithm is used in the propose system to improve the performance of search based face annotation scheme. The work of this paper is different from previous papers or existing systems; there are two key points that differentiate the proposed work from existing work. First key point is that it is used to solve general content based face annotation problem using search based face annotation scheme where face image is used as query image. And second key point is that unsupervised label refinement algorithm is used as suitable algorithm for enhancement of new label matrix for weakly labelled facial images. This work also related recent work of the WLRLCC method. The SBFA consists of following steps as shown in fig. 2.

Step 1: Facial image data collection;
Step 2: Face detection and facial feature extraction;
Step 3: Facial feature indexing;
Step 4: Refine weakly labeled data;
Step 5: Similar face retrieval.

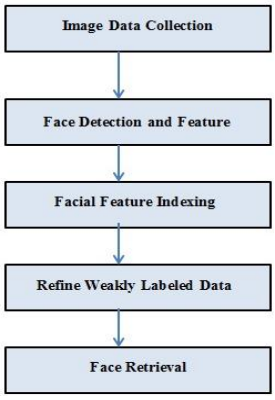

Fig 2: Search Based Face Annotation steps

E. Purifying web facial images

M. Zhao et al. [14] proposed a system that can learn and recognize faces by combining signals from large scale weakly labelled text, image, and video corpora. First, consistency learning is proposed to create face models for popular persons. It uses the text-image co-occurrence on the web as a weak signal of relevance and learns the set of consistent face models from this very large and noisy training set. To recognize people in videos, face detection and tracking are applied to extract faces from videos. Then, key faces are selected for each track for fast and robust recognition. Face tracks are further clustered to get more compact and robust representation. The face tracks are further clustered to get more representative key faces and remove duplicate key faces. A combination of majority voting and probabilistic voting algorithm is used to recognize each cluster of face tracks. They showed various active learning possibilities in case of improving the recognizer to grow across age variations. Their proposed work provides another direction which would be shows that how to combine high precision face-based retrieval scheme and high-recall text-based retrieval scheme.

D. Wang et al. [15] this proposed work adopted a unified framework of Unifying Transductive and Inductive Learning (UTIL) for mining web facial images to tackle the face annotation problem by combining the strengths of the two learning techniques. They worked on Weak label Laplacian support vector machine (WL-LapSVM) algorithm by adopting WLRLCC algorithm. Proposed system has an entropy-based rank level fusion algorithm, which performs and supervise regression-based fusion algorithm without any extra training efforts.

III. APPLICATIONS

Face annotation can be used in various applications fields are as follows

- Identity verification (electoral registration, passports, drivers' licenses, employee IDs),
- Criminal justice systems (forensics),
- Wild landmark face annotation,
- Online photo album management, and
- Social media sites like Facebook (In case of Facebook face annotation can be termed as 'tagging'.)

IV. CONCLUSIONS

This paper presents a survey on face annotation techniques for web facial images. Currently, many new methods are proposed in the field of auto face annotation. This paper shows that many research problems have been highlighted and direction for future work has been suggested in existing works. Many open issues have been highlighted by the researchers such as dealing with auto face annotation on large scale databases by different techniques. From overall study of this paper suggests that use of unsupervised labelled refinement (ULR) approach with clustering based approximation algorithm improves efficiency and scalability of search based scheme. Evaluation of auto face annotation on large scale database is depending on voting schemes implementing for system.

ACKNOWLEDGMENT

I would like to express my special appreciation and thanks to my advisor Professor Mrs. Archana Chaugule, you have been a tremendous mentor for me.

REFERENCES

[1] L. Zhang, L. Chen, M. Li, and H. Zhang, "Automated Annotation of Human Faces in Family Albums," Proc. 11th ACM Int'l Conf. Multimedia (Multimedia), 2003.

[2] P.T. Pham, T. Tuytelaars, and M.-F. Moens, "Naming People in News Videos with Label Propagation," IEEE Multimedia, vol. 18, no. 3, pp. 44-55, Mar. 2011.

[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] G.B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments," technical report 07-49, 2007.

[6] 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.

[7] J. Fan, Y. Gao, and H. Luo, "Hierarchical Classification for Automatic Image Annotation," Proc. 30th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), pp. 111-118, 2007.

[8] 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.

[9] J. Tang, R. Hong, S. Yan, T.-S. Chua, G.-J. Qi, and R. Jain, "Image Annotation by KNN-Sparse Graph-Based Label Propagation over Noisily Tagged Web Images," ACM Trans. Intelligent Systems and Technology, vol. 2, pp. 14:1-14:15, 2011.

[10] J.Y. Choi, W.D. Neve, K.N. Plataniotis, and Y.M. Ro, "Collaborative Face Recognition for Improved Face Annotation in Personal Photo Collections Shared on Online Social Networks," IEEE Trans. Multimedia, vol. 13, no. 1, pp. 14-28, Feb. 2011.

[11] M. Guillaumin, T. Mensink, J. Verbeek, and C. Schmid, "Automatic Face Naming with Caption-Based Supervision," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2008.

[12] M. Guillaumin, T. Mensink, J. Verbeek, and C. Schmid, "Face Recognition from Caption-Based Supervision," Int'l J. Computer Vision, vol. 96, pp. 64-82, 2011.

[13] Z. Wu, Q. Ke, J. Sun, and H.-Y. Shum, "Scalable Face Image Retrieval with Identity-Based Quantization and Multi-Reference Re-Ranking," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 3469-3476, 2010.

[14] M. Zhao, J. Yagnik, H. Adam, and D. Bau, "Large Scale Learning and Recognition of Faces in Web Videos," Proc. IEEE Eighth Int'l Conf. Automatic Face and Gesture Recognition (FG), pp. 1-7, 2008.

[15] D. Wang, S.C.H. Hoi, and Y. He, "A Unified Learning Framework for Auto Face Annotation by Mining Web Facial Images," Proc. 21st ACM Int'l Conf. Information and Knowledge Management (CIKM), pp. 1392-1401, 2012.

[16] D. Wang, Steven C.H. Hoi, Member, IEEE, Ying He, and Jianke Zhu, "Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation," IEEE Transactions On Knowledge And Data Engineering, vol. 26, No. 1, January 2014.

