IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Special Issue-12, May 2019 E-ISSN: 2347-2693

Fake Data Mining over Distributed Database With Face Annotation

Pankaj S Wankhede[1*], Sachin Choudhari[2], Ashish Kumbhare[3]

[1,2,3]M.Tech, Department of Computer Science & Engineering, Jhulelal Institute of Technology, Nagpur, India

Corresponding Author: pankajwankhede5@gmail.com

Available online at: www.ijcseonline.org

***Abstract*—** A face annotation has many applications the main part of based face annotation is to management of most same facial images and their weak data labels. This problem different method are adopted. The efficiency of annotating systems are improved by using these methods. This paper proposes a review on various techniques used for detection and analysis of each technique. Combine techniques are used in retrieving facial images based on query. So it is effective to label the images with their exact names. The detected face recognition techniques can annotate the faces with exact data labels which will help to improve the detection more efficiently. For a set of semantically similar images Annotations from them. Then content-based search is performed on this set to retrieve visually similar images, annotations are mined from the data descriptions. The method is to find the face data association in images with data label. Specifically, the task of face-name association should obey the constraint face can be a data appearing in its associated a name can be given to at most one face and a face can be assigned to one name.

***Keywords*—** *Annotation, weak data, exact data, detection, Content Based.*

I. INTRODUCTION

The face annotation is an important technique that to annotate facial feature images automatically. The face annotation can be useful to many Applications. The face annotation approaches are often treated as an extended face recognition issue, where different classification models are trained model-based face annotation time consuming and cost of to collect a large amount of human labelled facial images. some studies have attempted to get a search based annotation for facial image annotation by mining to tackle the automated face annotation by exploiting content-based image retrieval method.[1] The objective of is to assign correct data labels given query facial image. It is usually time consuming and cost to collect a large amount of human data labeled training facial images. It is usually difficult to the models when new data or new persons are added, in which an retraining process is usually required. The annotation or recognition performance often poorly when the number of persons or classes is very large[2]. In, face annotation was conducted by users on image viewer mode. That is, if a user desires to label a face, he/she moves mouse onto that face, and a candidate name list will popup to provide a recommendation. The user can either select a name from the name list, or set a new name for that face. Algorithm wise, the system tries to generate the candidate name list from historical labeling results. In contrast, in the proposed new approach, the browsing mode is supported for face annotation and name propagation. A user can select a number of photographs in the browsing mode and then perform a one-click assignment of a name to all of those photographs[3]. Algorithm wise, the system tries to infer the correspondence between name and face, i.e. propagate the name from image level to face level. Traditionally, thumbnails are generated by directly down sampling the original images. However, thumbnails generated by this approach are difficult to recognize, especially when the thumbnails are very small. To overcome such limitation, we employ a technique called "smart thumbnail" to facilitate browsing experiences[4]. With the detected faces and attention areas, the system automatically generates smart thumbnails. Initially, the system display is traditional down sampled thumbnails in browsing mode. When a user moves mouse onto an image, the system automatically switches the thumbnail to a smart thumbnail, displaying the most informative part of the image, i.e. face areas, to the user. If the mouse pointer hovers on the image for few seconds, the system will automatically animate the image with an optimal browsing path to let users browse through the different parts of an image[5]. In this way, users will rarely suffer from recognizing individuals from small thumbnails. With viewer mode and browsing mode, face annotation is conducted naturally and efficiently. However, the process could be more efficient if we integrate the two, because users may frequently alternate between viewer mode and browsing mode. For name list generation algorithm in viewer mode, there have some additional inputs from browsing mode, i.e. dangling names on images. For name propagation algorithm in browsing mode, there also have some additional inputs from viewer mode, i.e. labeled faces in some images. Taking into account these additional inputs, face annotation will be more efficient. In addition, given the face similarity measure function, the system also provides the function of similar face retrieval. In this way, users are allowed to search similar faces by specifying either a face or a name and then annotate multiple faces in a batch way. 3.2 Face Representation As current face recognition algorithms are not robust enough for face annotation in family album systems, in the proposed face annotation framework, contextual features are incorporated in addition to those used in classical face recognition algorithms, as proposed. Therefore, each face will be represented by both facial appearance feature and contextual feature. Following the same approach, facial appearance features are extracted based on face detection result. Because face appearance features are most reliable when extracted from frontal faces, a texture-constrained active shape model is applied to determine if an input face is in frontal view. If the face is not in frontal, the face appearance feature is treated as a missing feature. After this process, each face is geometrically aligned into the standard normalized form to remove the variations in translation, scale, in-plane rotation and slight out-of-plane rotation. Then facial appearance features are extracted from normalized gray face images. The contextual features are extracted from the extended face region. Compared to contextual features used, we add color moment feature in LUV color space to compensate the lack of global color feature. By dividing the extended face region into 2×1 blocks, local regional features are extracted to capture the structural information of body patches. We also restrict that the date difference between two photographs must be within two days when comparing their contextual similarity.

II. Literature Review

This investigates a 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 search-based face annotation scheme is how to effectively perform annotation by exploiting the list of most similar facial images and their weak labels that are often noisy and incomplete. To tackle this problem, we propose an effective unsupervised label refinement (ULR) approach for refining the labels of web facial images using machine learning techniques. We formulate the 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 the scalability considerably. We have conducted an extensive set of empirical studies on a large-scale web facial image testbed, in which encouraging results showed that the proposed ULR algorithms can significantly boost the performance of the promising SBFA scheme[1].The face annotation has many real world applications. The challenging part of search based face annotation task is management of most familiar facial images and their weak labels. To tackle this problem, different techniques are adopted. The efficiency and performance of annotating systems are improved tremendously by using these methods. Here this paper proposes a review on different techniques used for this purpose and check the pros and cons of each technique[2].

Face images that are captured by surveillance cameras usually have a very low resolution, which significantly limits the performance of face recognition systems. In the past, super-resolution techniques have been proposed to increase the resolution by combining information from multiple images. These techniques use super-resolution as a preprocessing step to obtain a high-resolution image that is later passed to a face recognition system[3]. Considering that most state-of-the-art face recognition systems use an initial dimensionality reduction method, we propose to transfer the super-resolution reconstruction from pixel domain to a lower dimensional face space. Such an approach has the advantage of a significant decrease in the computational complexity of the super-resolution reconstruction[4]. The reconstruction algorithm no longer tries to obtain a visually improved high-quality image, but instead constructs the information required by the recognition system directly in the low dimensional domain without any unnecessary overhead. In addition, we show that face-space super-resolution is more robust to registration errors and noise than pixel-domain super-resolution because of the addition of model-based constraints[5].

III. Research Methodology

- Database creation with image in binary bit format array
- Scanning BMP Format Reading per pixel value in RGB value
- Facial feature indexing with data label
- Similar face retrieval with value
- Detected Final output
- Refined data

Data Labelling

Data labeling procedure. The procedure are compared with data labeling on spectral clustering. After initial labeling with partial clustering, The proposed labeling algorithm and spectral clustering to label the rest of the faces. We recluster label faces, then data label the cluster, which similarity variation is the lowest. proposed data labeling algorithm get higher efficiency at the beginning of data labeling, The selection of neural network is done as it has got the unique feature of flexibility with accuracy. We can consider neural network as that student who once taught a thing never forgets to reproduce it as and when required. FACE FEATURE RECOGNITION has adopted the ADAPTIVE behavior of neural network, which makes the project consider Time

Complexity and Space Complexity Human beings are gifted with a unique power of visualizing and interpreting the things. But this power is being misused sometimes or it may lead you in a state of ambiguity. It happens sometimes that criminals do change their get up to escape .we familiar with sort of event in some hijacking case or bank robbery case. But if the power of visualizing is adopted in system, the system will not allow this to happen comes for help here. FACE FEATURE RECOGNITION has neural net has learning element. A face is given as input to the neural net, it matches it with faces in its databases and give output in the form as it has RECOGNISE the face and will display the personality description. This is the core of the thesis. While giving the face as input few initial arrangements has to be carried out. These arrangements are listed below.

a) EDGE DETECTION: -

Def: edge detection is to detect the edges from the given figures. There are various methods available to achieve edge selection. These can be applied as per the available photograph.

TRACING INPUT PHOTOGRAPH:-

This is the method where by we traces the outline using a trace paper and a transparent paper. We take a trace paper and adjust lift over the photograph then using a sharp pencil, and good lighting, we trace out the outline of the photograph on the trace paper. Similarly, we take out 3-4 traces till we get trace as per the actual photograph. This face is then forwarded to the next module after thinning.

b) SCANNING TOP-DOWN AND LEFT-RIGHT:-

This is totally software approach. Here, we have the scanned image in BMP format inside profile. The first job to skip the header. Now, the pointer points to pixel info. This file is now scanned left to right and RGB values are extracted. If RGB value of two consecutive pixels is found to have difference greater than a given value, it is taken into accounts, and is displayed on screen. Otherwise, we move to scan next pixel in sequence without displaying anything. Once we have scanned in left to right fashion, now to scan in top to bottom fashion. Here, again RGB values of pixels are extracted. If the difference is above taken value, the pixel is displayed at its x-y co-ordinate position on screen; else we neglect the pixel. This is continued till whole image is scanned.

IV. RESULTS AND DISCUSSION

FILTERING: -

DEF: filtering means to filter out required information from the available figure leaving best the irregularities and scattered pixels.

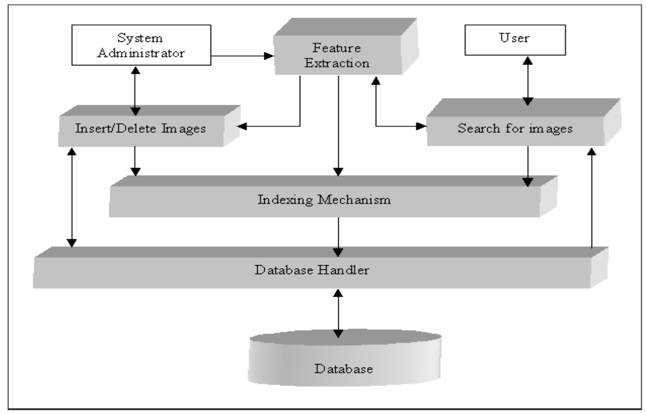

Figure 1. Data flow diagram

FACIAL FEATURES: -

a) At a Glance: -

This phase is one of the most vital phase of the project. The outline profile of the end view of human face pattern yields the best descriminant features for identification of human face. This involves the following steps: -

1) Binarizing the digitized photograph.
2) Reduction of the binarized photograph to line profile.

The ultimate output of this phase is ten facial features, which is necessary and enough to make neural network recognize the human face.

Selection of the Fiducially marks on the profile: -

If we look at the side profile of the human face, we find that certain points can be readily defined on the face profile. If these points can be correctly identified they can help in extracting certain characteristics features for that particular face. Ten such points are shown in figure .Out of these ten points, eight points are independent of each other but point 3 and 2 are interrelated with each other. All these points are calculated by using some mathematical relationship logic along with some statistical knowledge

Now, this image can be used to extract the ten points. The ten points and the methods to extract them is as follows:

A) Nose Point (Point1): -This point is the most important of all the points. Nose point forms the basis for the computation of all the other fiducial points. The other can be calculated using certain mathematical logic and available formulas only after finding nose points. To find the nose point the logic to be applied is that, the first point obtained as we scan left to right, the input photograph in side profile is considered as the nose point.

B) Chin Point (Point2): -this is the second fiducial point, which is to be calculated after the calculation of nose point. To find the chin point, all the pixels below the nose point, in the side profile of the input human face are joined with the nose point and whichever pixel is making the largest angle is being considered as the chin point.

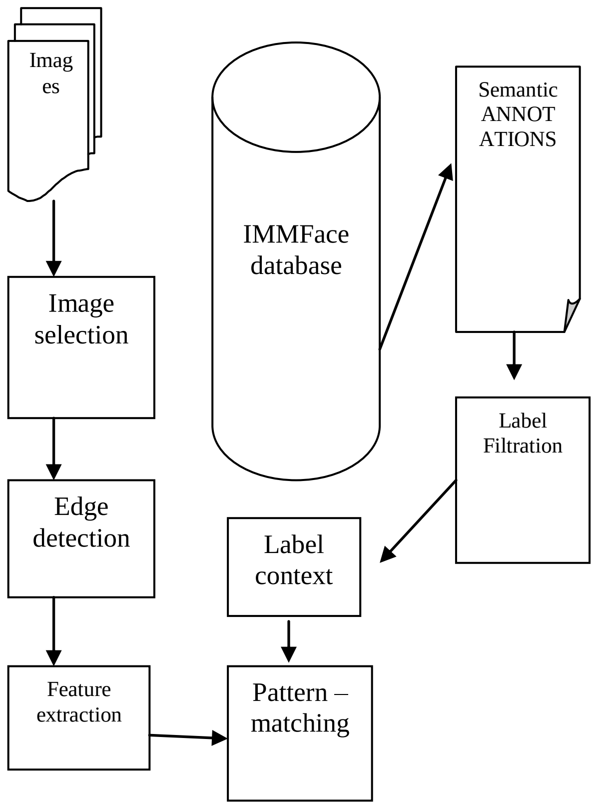

Figure 2. Flow Diagram

V. Conclusion and Future Scope

The face annotation on labeled images. So research works and new methods are being proposed. The research in this field importance as it is very useful in searching and social Media. The future work will work on multi person data task and thereby efficiency and accuracy of result. If the techniques are implemented properly, then the data label problem will be solved.

References

[1] Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation Dayong Wang, Steven C.H. Hoi, Member, IEEE, Ying He,
and Jianke Zhu JANUARY 2017.

[2] Dayong Wang, Steven C.H. Hoi, 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

[3] D. Wang, S.C.H. Hoi, Y. He, and J. Zhu, "*Retrieval-Based Face Annotation by Weak Label Regularized Local Coordinate Coding,*" Proc. 19th ACM Int'l Conf. Multimedia (Multimedia), pp. 353-362, 2011.

[4] W. Dong, Z. Wang, W. Josephson, M. Charikar, and K. Li, "*Modeling LSH for Performance Tuning*," Proc. 17th ACM Conf. Information and Knowledge Management (CIKM), pp. 669-678, 2008

[5] C. Siagian and L. Itti, "*Rapid Biologically-Inspired Scene Classification Using Features Shared with Visual Attention*,"IEEE Trans. Pattern Analysis and Machine Intelligence, vol.29, no. 2,pp. 300-312, Feb. 2007.

[6] 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.

[7] 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.

[8] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "*Face Recognition: A Literature Survey*," ACM Computing Survey, vol. 35, 2003.

[9] 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.

