Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 5.258

IJCSMC, Vol. 5, Issue. 3, March 2016, pg.322 – 329

EXTRACTING RELATIVE FACE NAMING SIMILARITY ON LOW SUPERIORITY IMAGE USING MIL

A.Mohamed Igbal
Dept. of CSE
IFET College of Engineering

Mrs. S.Usharani, SAP
Dept. of CSE
IFET College of Engineering

ABSTRACT: Labelling faces in the images is tremendously challenging due to huge variation in image appearance of each character and the weakness, ambiguity of available annotation. Some of the pre-processing steps to be carried out before performing face naming. Even though successfully performing pre-processing step such as face detector and entity name detector, face naming is still a challenging task. So, the supervised approach requires sufficient labelled training data in order to achieve high accuracy. The supervised learning method such as Multiple Instance Learning algorithm would solve the ambiguity of the labelling process by making weaker assumptions about the labelling information. Based on Multiple Instance Learning algorithm, the two discriminative probabilistic learning method, "Quasi-Positive Bags" and "Extended Diverse Density" is used to develop an automatic training scheme. This algorithm will be applicable for large image databases and images does not have complete data labels.

KEYWORDS: Face naming, Multiple Instance Learning, Quasi Positive Bags.

1. INTRODUCTION

In this paper, we focus on automatically annotating faces in images based on the ambiguous supervision from the associated captions. Some pre-processing steps need to be conducted before performing face naming. Specifically, faces in the images are automatically detected using face detectors, and names in the captions are automatically extracted using a name entity detector. Here, the list of names appearing in a caption is denoted as the candidate name set. Even after successfully performing these pre-processing steps, automatic face naming is still a challenging task. The faces from the same subject may have different appearances because of the variations in poses, illuminations, and expressions. Moreover, the candidate name set may be noisy and incomplete, so a name may be mentioned in the caption, but the corresponding face may not appear in the image, and the correct name for a face in the image may not appear in the corresponding caption. Each detected face (including falsely detected ones) in an image can only be annotated using one of the names in the candidate name set or as null, which indicates that the ground-truth name does not appear in the caption.

Recently, there is an increasing research interest in developing automatic techniques for face naming in images as well as in images. To tag faces in news photos, proposed to cluster the faces in the news images. A graph-based method by constructing the similarity graph of faces and finding the densest component. The multiple-instance logistic discriminant metric learning (MildML) method and proposed a structural support vector machine (SVM)-like algorithm called maximum margin set (MMS) to solve the face naming problem. A regularized low-rank representation (rLRR) by incorporating weakly supervised information into the low-rank representation (LRR) method, so that the affinity matrix can be obtained from the resultant reconstruction coefficient matrix. This all above the existing system are taking the too much time to detect the face.

Learning classifiers from the ambiguously labeled data falls in the category of ambiguous learning. The ambiguous association between samples and labels make the learning task more challenging than that in standard supervised learning.

Here, Multiple-Instance Learning (MIL) was proposed for machine learning to solve the ambiguity of the labeling process by making weaker assumptions about the labeling information. In this learning scheme, instead of giving the learner labels for individual examples, the trainer only labels collections of examples, which are called bags. A bag is labeled negative if all the examples in it are negative. It is labeled positive if there is at least one positive example in it. The key challenge in MIL is to cope with the ambiguity of not knowing which instances in a positive bag are actual positive and which are not. Based on that, the learner attempts to find the desired concept.

In this paper, we proposed a cross-modality automatic training method based on the MIL approach to recognize faces from a large news image dataset. In order to apply the MIL approach to the name-face association problem mentioned above, we introduce a new concept “Quasi-Positive bags” and propose a new approach “Extended Diverse Density” to handle the Quasi-Positive bags. The overall process of multi-modality training for name-face association.

2. RELATED WORK

Learning visual classifiers from caption-accompanying images has been an active topic in computer vision of which learning face classifiers from such da-ta is of particular interest. There are a few methods that explicitly take face-name (sample-label) correspondences into account. For example, Berg et al, pro-posed a constrained mixture model to optimize the likelihood of particular face-name assignment. The work in first iteratively clusters faces using EM based on face similarity and constraints from the caption. Based on these clusters, a weighted bipartite graph modelling the null assignment (i.e., faces that are not assigned to any names and names that are not assigned to any faces) and caption constraints is constructed for face-name assignment. On the other hand, Support Vector Machine (SVM) based methods directly learn discriminant classifiers using the ambiguously labeled data. Cour et al. proposed a max-margin formulation by introducing an ambiguous 0/1 loss to replace the loss in the standard SVM formulation, in which they defined the ambiguous 0/1 loss as 0 if the predicted name is in the image caption, and 1 otherwise. Based on this ambiguous loss, they defined a convex loss that penalized the prediction of names as the ones not present in the caption.

This formulation did not consider the uniqueness constraint, hence it generally cannot perform well for images with multiple faces. Luo et al. extended the idea of ambiguous loss for images with multiple faces, in which they enforced the uniqueness constraint by assigning names to faces at a set level (via labeling vectors) in each image.

Recently, low-rank property of a set of linearly correlated images shows its usefulness in many computer vision problems, such as subspace segmentation, face recognition, multi-label image classification, image alignment and image segmentation. On the other hand, PPM has been popularly used for feature point correspondence with unsupervised learning.

3. PROPOSED WORK

On automatically annotating faces in images based on the ambiguous supervision from the associated captions. So, proposed a cross-modality automatic training method based on the MIL approach to recognize faces from a dataset. Multiple-Instance Learning (MIL) proposed for machine learning to solve the ambiguity of the labelling process by making weaker assumptions about the labelling information. In order to apply the MIL approach to the name-face association problem, we introduce a new concept "Quasi-Positive bags" and propose a new approach "Extended Diverse Density" to handle the Quasi-Positive bags. The key challenge in MIL is to cope with the ambiguity of not knowing which instances in a positive bag are actual positive and which are not. Based on that, the learner attempts to find the desired concept. It will Achieve high quality of face naming detection scheme. It will analyze the ambiguity of instance. To examine the distribution of the instance vectors, and look for a feature vector that is close to the instances from different positive bags and far from all the instances from the negative bags.

4. SYSTEM MODEL

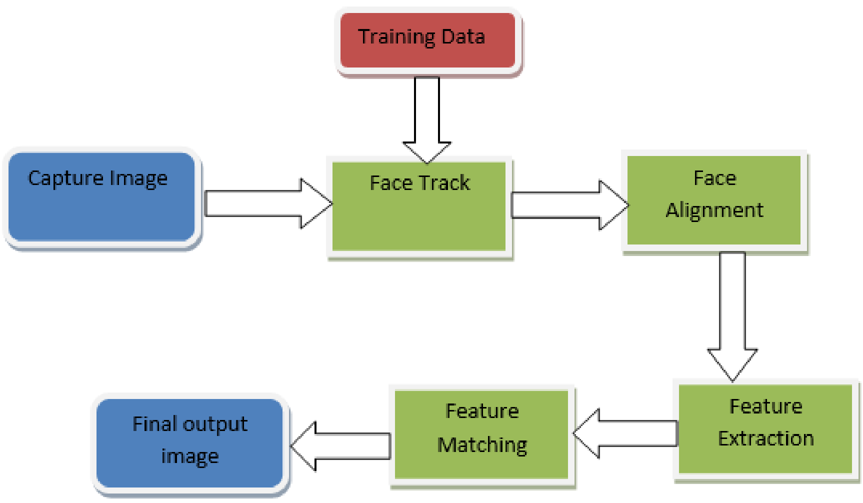


Fig 1 : System Architecture

4.1. Input Image

In this module, we first taken the image as input and an input image is further taken into process for creating training dataset. Based on training image dataset and given image set, it will start to analyze the feature extraction from the input and training set images.

4.2. Feature generation

To generate the features needed for the algorithm, images are separated into shots with one keyframe for each shot. Also, we just focus on the frontal face model. We first extract frontal faces from keyframes, use skin detection to exclude some false alarm detections, and then obtain the eigenfaces for face recognition.

4.2.1 Feature extraction on Skin Detection

Our skin detection algorithm is based on a skin pixel classifier, which is derived using the standard likelihood ratio approach. Let x represents a vector containing the color values at a pixel. Then x is labeled as skin, we set the threshold as 1. (| skin) Px and (|not skin) Px are estimated as Gaussian models with a diagonal covariance matrix with a size of 28 x28 .

After getting skin pixel candidates, we post-process the candidates to determine the skin regions, using techniques including Gaussian blurring, thresholding, and mathematical morphological operations such as closing and opening.

4.2.2 Feature extraction on Eigen face generation

The frontal faces, which are in a relatively large scale (larger than 48 48 ×), and include certain skin regions (which cover more than a quarter of the face region obtained from the "Haar-like face detector" algorithm), are detected from keyframes in the image sequences. After normalized to a size of 64 64 × , multiplied by a constant to map the median value of the pixels to 128, and limited the pixel values to 255, they are used to get the top 22 eigenfaces with 85% energy for recognition. The features used throughout this are the projection coefficients based on these eigenfaces.

4.3. Cross-modality automatic training

This is how to find the Quasi-Positive bags and the negative bags for learning the model based on MIL.

4.3.1 Quasi-Positive bag generation

The Quasi-Positive bags are those frames which are associated with the names mentioned in the image data. When a person will appear in the following images. Therefore, our algorithm automatically selects candidate frames, which are believed to be with high probability to have the face of that person, according to the association between sequence of images. Here, Closed Captions (CC) data as the candidate frames because those four frames have this face with a relatively high probability based on our observation.

4.3.2 Negative bag generation

Our objective is to find a common point from all the Quasi-Positive bags. The useful negative instances are those confusing negative examples in the Quasi-Positive bags, such as the interested persons.

5. Multiple-Instance Learning

In this section, we present a brief introduction to MIL and Diverse Density.

5.1 Multiple-Instance Learning problem

The MIL model was first formalized by Dietterich et al. [7] to deal with the drug activity prediction problem. Following that, an algorithm called Diverse Density was developed in [8] to provide a solution to MIL, which performs well on a variety of problems such as drug activity prediction, stock selection, and image retrieval.

More formally, given a set of instances x1,x2,...,xN, the task in a typical machine learning problem is to learn a function.

$$y = f(x_1, x_2, \cdots, x_N)$$

so that the function can be used to classify the data. In traditional supervised learning, some training data are given in terms of (yi,xi). Based on those training data, the function is learned and used to classify the data outside the training set.

In MIL, the training data are grouped into Bags, with $X_j = \{\mathbf{x}_i : i \in I_j\}$ and $I_j \subseteq \{1, \cdots K\}$

Instead of giving the labels yi for each instance, we have the labels Yi for each bag. A bag is labeled negative (Yj =−1), if all the instances in it are negative. A bag is positive (Yj= 1), if at least one instance in it is positive.

5.2 Diverse Density

One way to solve MIL is to examine the distribution of the instance vectors, and look for a feature vector that is close to the instances from different positive bags and far from all the instances from the negative bags. Such a vector represents the concept we are trying to learn. This is the basic idea of the Diverse Density (DD) algorithm. Diverse Density is a measure of the intersection of the positive bags minus the union of the negative bags. By maximizing Diverse Density, we can find the point of intersection (the desired concept). Here a simple probabilistic measure of Diverse Density is explained.

5.3 Extended Diverse Density

In our application, what we have are Quasi-Positive bags, i.e., some positive bags do not include positive instances at all (which are called false-positive bags in this paper). In a false-positive bag, by the original DD definition, Pr (t|Bi+) will be very small or even zero. These outliers will influence the DD significantly due to the multiplication of the probabilities.

To avoid the influence of false-positive bags, we propose "Extended Diverse Density" (EDD) to handle the Quasi-Positive bag problem for MIL. where Pr (t|Bi+) and Pr (t|Bi-) are also estimated by equation (3), and Zis a normalization parameter which is the total number of instances in the Quasi-Positive bags, which makes the value within the range of [0,1].

By using summation instead of multiplication, the false-positive bags which possess smaller Pr (t|Bi+) will make little influence to the EDD value, when tis the true concept, while the truly positive bags will still contribute large values to the EDD.

The EDD is defined as:

$$\arg\max_t \frac{\sum_i \Pr\left(t \mid B_i^+\right)}{Z} \prod_i \Pr\left(t \mid B_i^-\right)$$

5.4 Face description

Face images are extracted using the bounding box of the Viola-Jones detector and aligned using the funneling method (Huang et al., 2007a) of the Labeled Faces in the Wild data set. This alignment procedure finds an affine transformation of the face images so as to minimize the entropy of the image stack. On these aligned faces, we apply a facial feature detector (Everingham et al., 2006). The facial feature detector locates nine points on the face using an appearance-based model regularized with a tree-like constellation model.

There is a large variety of face descriptors proposed in the literature. This includes approaches that extract features based on Gabor filters or local binary patterns. Our work in Guillaumin et al. (2009b) showed that our descriptor performs similarly to recent optimized variants of LBP for face recognition (Wolf et al., 2008) when using standard distances. Our features are available with the data set.

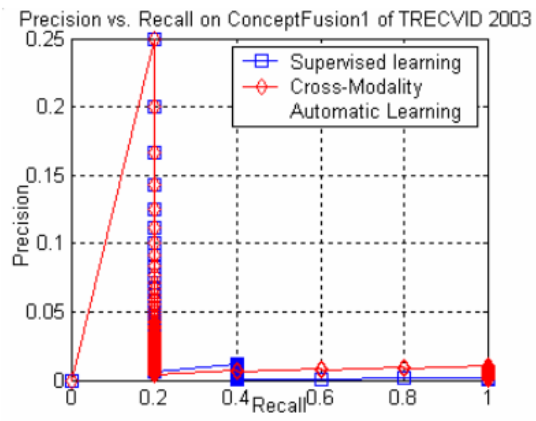
5.5 Metrics for face identification

For both the face retrieval tasks and the face naming tasks, we indeed need to assess the similarity between two faces with respect to the identity of the depicted person. Intuitively, this means that a good metric for identification should produce small distances – or higher similarity – between face images of the same individual, while yielding higher values – or lower similarity – for different people. The metric should suppress differences due to pose, expression, lighting conditions, clothes, hair style, sun glasses while retaining the information relevant to identity. These metrics can be designed in an ad-hoc fashion, set heuristically, or learned from manually annotated data.

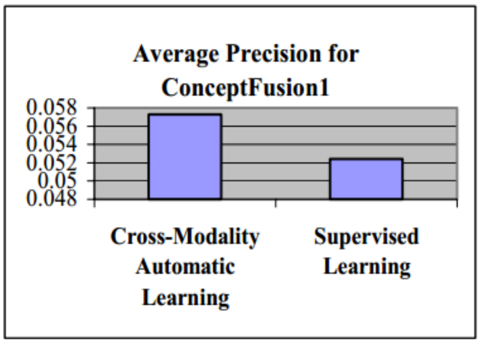
6. EXPERIMENTAL RESULTS

We now demonstrate the performance of our algorithm from the NIST Image TRECVID 2003 corpus. The whole image dataset is divided into five parts: Concept Training, ConceptFusion1, ConceptFusion2, Concept Validate, and Concept Testing. The first experiment is to train a frontal face model for "Bill Clinton" based on a image sequence "19980205_ABC". There are totally 9 Quasi-Positive bags generated, with 6 false-positive bags.

In the second experiment, a model for recognizing "Madeleine Albright" is trained in Concept Training, which includes 78 news images with 28,055 key frames. In this case, eight out of twenty four Quasi-Positive bags include "Madeleine Albright". Quasi-Positive bags with the detected faces marked by rectangles. In our experiment, the model for recognizing the specified person is learned as the maximum point of the EDD in the eigen space. The top eight closest faces to this model in the training data are showed, which include six correct answers. Among them, the closest face has the highest EDD of 0.2338. Quasi-Positive bags as well as the top three closest faces to the model of "Newt Gingrich".



(a) ROC curves



(b) Average Precision

Fig. 2 Performance comparison

7. CONCLUSIONS

We have presented a cross-modality automatic training algorithm based on Multiple-Instance Learning for face recognition in a large news image dataset. Based on prior information about news image, Quasi-Positive bags and negative bags are generated. Extended Diverse Density was proposed to handle the Quasi-Positive bags in order to find the concept we are interested in. We also propose to use the "Relative Sparsity" of a cluster to detect the anchor person in the news images. Our model is tested and compared with a supervised model in database ConceptFusion1, and shows a promising result. Ongoing works include trying to use more features, such as the skin tone and the shape for learning a common model; tracing the shot to find more suitable candidates for learning the recognition model; and perform more simulations to characterize the performance and compare the performance with other approaches.

REFERENCES:

[1] M. Turk and A. Pentland, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, Vol. 3, No. 1, pp. 71-86, Winter 1991.

[2] L. Wiskott, J.M. Fellous, N. Kr uger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 775-779, 1997.

[3] W. Zhao, R. Chellappa, and J. Phillips, “Subspace Linear Discriminant Analysis for Face Recognition,” Technical Report, CS-TR4009, Univ. of Maryland, 1999.

[4] W. Zhao, R. Chellappa, A. Rosenfeld, and P.J. Phillips. Face recognition: A literature survey. CVL Technical Report, University of Maryland, October 2000.

[5] S. Satoh and T. Kanade. “Name-it: Association of face and name in video,” Proc. of CVPR’97, Puerto Rico, 1997.

[6] O. Maron. “Learning from ambiguity,” PhD dissertation, Department of Electrical Engineering and Computer Science, MIT, Jun. 1998.

[7] T. G. Dietterich, R. H. Lathrop and T. Lozano-P_erez, “Solving the multiple instance problem with axis-parallel rectangles,” Artificial Intelligence Journal, 89, pp. 31-71, 1997.

[8] M. Everingham, J. Sivic, and A. Zisserman, “Hello! My name is...Buffy—Automatic naming of characters in TV video,” inProc. 17thBrit. Mach. Vis. Conf., Edinburgh, U.K., Sep. 2006, pp. 899–908.

[9] J. Sang and C. Xu, “Robust face-name graph matching for movie charac-ter identification,” IEEE Trans. Multimedia, vol. 14, no. 3, pp. 586–596,Jun. 2012.

[10] Y.-F. Zhang, C. Xu, H. Lu, and Y.-M. Huang, “Character identification in feature-length films using global face-name matching,” IEEE Trans. Multimedia, vol. 11, no. 7, pp. 1276–1288, Nov. 2009.

[11] X. Zhang, L. Zhang, X.-J. Wang, and H.-Y. Shum, “Finding celebrities in billions of web images,” *IEEE Trans. Multimedia*, vol. 14, no. 4, pp. 995–1007, Aug. 2012

[12] M. Özcan, L. Jie, V. Ferrari, and B. Caputo, “A large-scale database of images and captions for automatic face naming,” in *Proc. 22nd Brit. Mach. Vis. Conf.*, Dundee, U.K., Sep. 2011

