Abstract

Annotation is a metadata attached to text, image or other data. There are many annotation techniques available for tagging the facial images. As millions of images are shared by the users with some name tagged in it, but it is not done properly. The proposed Search-based face annotation framework is data driven and model free which solve the problem that exist by mining weakly labelled facial images those are freely available on the internet. The SBFA uses an algorithm named Adaboost which is one of the prominent algorithms in annotating the images. The algorithm produces more efficient performance than other algorithms as this comes under supervised machine learning techniques. The results prove that the algorithm solves the above said problem i.e. eliminating duplication and weakly labelled facial images.

1. Introduction

In recent years due to rapid growth of social networking tools for web based photo sharing have witnessed enormous number of web images were captured and stored by the users. Most commonly the images are human images and which is not tagged properly. Due to this, when the user query’s an input image it finds difficult to identify the image. This issue led to the study of the automated annotation which helps to annotate the images automatically [8], [9], [10], [11]. Automated face annotation benefits many real world applications, one such application is Facebook photo sharing. Classical annotations uses different models to collect the image and then employ
learning techniques to annotate the images which is often treated as an issue in recognizing an image. However, model based annotations also limited to several features. In recent times some studies were undergone to find an efficient search based face annotation by mining the World Wide Web where huge numbers of weakly labelled images are freely available. The Search Based Face Annotation provides solution to annotate the weakly labelled images automatically using Supervised Learning techniques. The main objective of SBFA is to allocate accurate labels to a given input query image. One main issue that SBFA [6] is how to scrutiny the web images and their tags for automatic annotation task. To overcome the above issue we intend a Search Based annotation scheme. In particular, we put forward a novel Supervised Label Refinement (SLR) by investigating machine learning techniques to bring out pure labels from set of weakly labelled web facial images. The main contribution of this work includes the following:

- We present a promising Search Based Annotation Scheme through mining huge weakly labelled images available on web.
- We present a novel SLR approach for enhancing label quality through Adaboost algorithm.
- We have shown encouraging results by conducting some set of experiments.

**a. Data Mining**

Data mining automates the detection of relevant patterns in a database, using defined approaches and algorithms to look into current and historical data that can then be analysed to predict future trends. Because data mining tools predict future trends and behaviours by reading through databases for hidden patterns, they allow organizations to make proactive, knowledge-driven decisions and answer questions that were previously too time-consuming to resolve.

Data mining is not particularly new – statisticians have used similar manual approaches to review data and provide business projections for many years. Changes in data mining techniques, however, have enabled organizations to collect, analyse, and access data in new ways. The first change occurred in the area of basic data collection. Before companies made the transition from ledgers and other paper-based records to computer-based systems, managers had to wait for staff to put the pieces together to know how well the business was performing or how current performance periods compared with previous periods. As companies started collecting and saving basic data in computers, they were able to start answering detailed questions quicker and with more ease.

Data mining is not all about the tools or database software that being used. The user can perform data mining with comparatively modest database systems and simple tools, including creating and writing user’s own, or using off the shelf software packages. Complex data mining benefits from the past experience and algorithms defined with existing software and packages, with certain tools gaining a greater affinity or reputation with different techniques.

**b. Face Annotation**

Automatic image annotation (also known as automatic image tagging or linguistic indexing) [8] is the process by which a computer system automatically assigns metadata in the form of captioning or keywords to a digital image. This application of computer vision techniques is used in image retrieval systems to organize and locate images of interest from a database.

This method can be regarded as a type of multi-class image classification with a very large number of classes - as large as the vocabulary size. Typically, image analysis in the form of extracted feature vectors and the training annotation words are used by machine learning techniques to attempt to
automatically apply annotations to new images. The first methods learned the correlations between image features and training annotations, then techniques were developed using machine translation to try to translate the textual vocabulary with the 'visual vocabulary', or clustered regions known as **blobs**. Work following these efforts has included classification approaches, relevance models and so on.

The advantages of automatic image annotation versus content-based image retrieval (CBIR) are that queries can be more naturally specified by the user. CBIR generally (at present) requires users to search by image concepts such as colour and texture, or finding example queries. Certain image features in example images may override the concept that the user is really focusing on. The traditional methods of image retrieval such as those used by libraries have relied on manually annotated images, which is expensive and time-consuming, especially given the large and constantly growing image databases in existence.

## 2. Related Work

The work is closely related to several research works which are as follows:

- The first related work is based on Face Recognition and Verification [11]. Recent years have seen some upcoming studies of unconstrained face detection and verification techniques on web images such as LWF studies. Comprehensive reviews can be found in some survey papers for this topic.

- The second related work is about Image Annotation [8]. Classical techniques apply existing image recognition approaches for training the models through trained images and resulting in probabilities or correlation between the images. Given limited training images semi supervised methods results with labelled and unlabelled data which leads to time consumption and well labelled training images an expensive one. Unlike these existing works we present a novel Supervised Label Refinement approach which focusses on optimizing the label quality for Search Based Face Annotation task.

- The third related work is about annotating the face [8]. Most studies have focussed on annotation of personal images which contain high contextual clues such as family names, social values, geotags and so on. These approaches produce fairly actual results which are successfully deployed in some famous applications such as Google Picasa, Apple iPhoto and Facebook auto tagging solution.

- The fourth related work is about Annotating Web Facial Images [8] through Mining. Most of the studies considers human name as input query and try to refine text based search [7] results through finding consistency in web facial images visually. Recently a discriminant technique was put forth to improve generative approaches and to eliminate explicit computation in graph based approaches [9]. For facial images, Berg et al. proposed a model combining with clustering technique to estimate the relationship between web facial images and names in the labels.

- Our work is different from the previous works from the two main aspects. Firstly, the proposed approach provides a solution to content based face annotation problem using SBFA [6] where web images are directly used as input query and corresponding labels are returned as the result. Secondly, based on initial weak labels the proposed Supervised Label Refinement technique learns an enhanced new label matrix for all web images in a single name space.
Finally, we note that our work is also related to recent work of ULR CDA method in [7] and our latest work on Unified Learning scheme in [5]. Instead of enhancing the label matrix and retrieving the top images is different from SLR task in this paper. We note that learning approaches for Supervised Label Refinement job are inspired by some studies in Machine Learning including Supervised Learning and Multi Label techniques.

2.1 Search Based Face Annotation

Fig 2.1 suggests the system flow of proposed architecture of Search Based Face Annotation [6] which includes the following processes:

1. Facial Image Data Collection.
2. Face Detection and Facial Feature Extraction.
3. Facial Feature Indexing.
5. Similar Face Retrieval.
6. Face Annotation by Majority Voting on Similar images with refined labels.

The first four steps are conducted before face annotation task testing period and the last two are conducted during testing period of image annotation.

The first step is collection web facial images from existing web search engine according to name list that contains names of persons. Given the nature of web facial images, they are noisy and do not relate to right human name. So we call them as weakly labelled web facial images.

The second is to pre-process these web facial images in order extract information related to the face, including face detection and alignment, facial region extraction and its feature representation. For face detection and alignment we adopt supervised face alignment approach. For facial feature representation we extract GIST texture features to represent extracted faces.

The third step is to index the extracted features of web faces through a high dimensional indexing method for facilitating the job of similar face retrieval in the subsequent step. In our approach, we
adopt an effective and high performance indexing approach called Locality Sensitive Hashing (LSH) [4], [6].
The key step of the put frothed framework is to attain utmost enhanced label quality of weakly labelled web facial images through Supervised Learning Scheme were label quality plays a vital role in final annotation processes.
Next, we describe the process of web facial image annotation during testing phase. We first conduct a similar face retrieval task through a given query facial image to search for a subset of most similar faces. With the set of top k similar faces then we annotate the web facial images with a label by employing a majority voting approach.

3. Supervised Learning Techniques

Supervised learning technique is a machine learning technique for inferring the function from labelled training data sets. The data set consists of training examples, where each example is a pair consisting of input objects and desired output value. A Supervised learning technique analyses the training data and results in inferring function which is used for mapping new examples. This requires learning approach to generalize training data in a reasonable way.

In order to solve a given problem of supervised learning, one has to perform the following steps:

I. Define the type of training samples. The user should decide what kind of data is to be used as a training sample.

II. Gather a training sample. The training sample needs to be representative of the real-world use of the function. Thus, a sample of input objects is collected and equivalent outputs are also collected.

III. Define the input feature representation of the learned function. The accuracy of the learned function depends strongly on how the input object is represented. The input object is transformed into a feature vector, which includes a number of features that are expressive of the object. The number of features [10] should not be too big, because of the curse of dimensionality; but should contain enough information to accurately predict the output.

IV. Define the structure of the learned function and equivalent learning algorithm.

V. Complete the design. Run the learning algorithm on the collected training sample. Some supervised learning algorithms require the user to decide certain control parameters. These parameters may be adjusted by optimizing performance on a subset (called a validation set) of the training samples, or through cross-validation [1].

VI. Calculate the accuracy of the learned function. After parameter adjustment and learning, the performance of the resulting function must be measured on a trial sample that is separate from the training sample.

3.1 Approaches and Algorithms

- Boosting (meta-algorithm) [2], [3]
- Bayesian statistics
- Decision tree learning
- Naïve bayes classifier
- Nearest Neighbor Algorithm
- Support vector machines
- Random Forests
- Ensembles of Classifiers
3.2 Issues to be Considered
- Bias variance trade-off
- Function complexity and amount of training data
- Dimensionality of the input space
- Noise in the output values

3.3 Boosting (meta-algorithm)

Boosting is a machine learning ensemble meta-algorithm [2] for decreasing bias primarily and also variance in supervised learning, and a family of machine learning algorithms which convert weak learners to strong ones. While boosting is not algorithmically controlled, most boosting algorithms consist of iteratively learning weak classifiers with respect to a distribution and adding them to a final strong classifier. When they are added, they are characteristically weighted in some way that is usually associated to the weak learners’ correctness. After a weak learner is added, the data is reweighted: examples that are misclassified gain weight and examples that are classified properly lose weight.

The main variation between many boosting algorithms is their method of weighting training data points and hypotheses. AdaBoost is very popular and perhaps the most noteworthy historically as it was the first algorithm that could adapt to the weak learners. However, there are many more recent algorithms such as LPBoost, TotalBoost, BrownBoost, MadaBoost, LogitBoost, and others. Many boosting algorithms fit into the AnyBoost framework, which shows that boosting does slope descent in function space using a convex cost function.

3.3.1 AdaBoost

AdaBoost can be used in concurrence with many other learning algorithms to develop their recital. The output of the other learning algorithms ('weak learners') is collected into a weighted sum that characterizes the final output of the boosted classifier. AdaBoost is adaptive such that successive weak learners are squeezed in favor of those instances misclassified by previous classifiers. AdaBoost is sensitive to noisy data and outliers. In some problems, however, it can be less vulnerable to the overfitting problem than other learning algorithms. The individual learners can be weak, but as long as the performance of each one is slightly better than random guessing the final model can be established to meet to a strong learner. While every learning algorithm will lean towards to ensemble some problem types better than others, and will have many different parameters and configurations to be adjusted before achieving optimal performance on a dataset, AdaBoost is often referred to as the best out-of-the-box classifier.

When used with decision tree learning, information collected at each stage of the AdaBoost algorithm about the relative 'hardness' of each training sample is nurtured into the tree growing algorithm such that later trees tend to focus on harder to classify examples.

3.3.1.1 Algorithm

Training

AdaBoost refers to a particular method of training a boosted classifier [2]. A boost classifier is a classifier in the form

\[ F_T(x) = \sum_{t=1}^{T} f_t(x) \] (1)
where each \( f_t \) is a weak learner which takes an object \( x \) as input and results in real valued indicating the class of the object. The sign of the weak learner output recognizes the foreseen object class and the absolute value gives the assurance in that classification. Similarly, the \( T \)-layer classifier will be positive if the sample is in positive class else negative.

Each weak learner produces an output, hypothesis \( h_t(x_i) \), for each sample in the training sample. At each iteration \( t \), a weak learner is selected and assigned a coefficient \( \alpha_t \) such that the sum training error \( E_t \) of the resulting \( t \)-stage boost classifier is minimized.

\[
E_t = \sum_{i} E[F_{t-1}(x_i) + \alpha_t h(x_i)]
\]

Here \( F_{t-1}(x) \) is the boosted classifier that has been built up to the previous stage of training, \( E[F] \) is some error function and \( f_t(x) = \alpha_t h(x) \) is the weak learner that is being considered for addition to the final classifier.

**Weighting**

In each iteration of the training process, a weight is assigned to each sample in the training sample equal to the current error \( E(F_{t-1}(x_i)) \) on that sample. These weights can be used to inform the training of the weak learner, for instance, decision trees can be grown that favor splitting sets of samples with high weights.

**Algorithm: Supervised Learning Refinement - Adaboost**

**Input:** Samples \( x_1 \ldots x_n \)

**Output:** \( y_1 \ldots y_n, y \in \{-1, 1\} \)

**Initial weights** \( w_{1,1} \ldots w_{n,1} \) set to \( \frac{1}{n} \)

**Error function** \( E(f(x), y, i) = e^{-y_i f(x_i)} \)

**Weak learners** \( h: x \rightarrow [-1, 1] \)

For \( t \) in \( 1 \ldots T \)

- Choose \( f_t(x) \):
  - Find weak learner \( h_t(x) \) that minimizes \( \epsilon_t \), the weighted sum error for misclassified points
    \[
    \epsilon_t = \sum_{i} w_{i,t} E(h_t(x), y, i)
    \]
  - Choose
    \[
    \alpha_t = \frac{1}{2} \ln \left( \frac{1 - \epsilon_t}{\epsilon_t} \right)
    \]
Add to ensemble:
\[ F_t(x) = F_{t-1}(x) + \alpha_t h_t(x) \]

Update weights:
\[ w_{i,t+1} = w_{i,t} e^{-y_i \alpha_t h_t(x_i)} \text{ for all } i \]
\[ \sum w_{i,t+1} = 1 \]

Renormalize \( w_{i,t+1} \) such that
\[ \frac{\sum_{h_t(x_i) = y_i} w_{i,t+1}}{\sum_{h_t(x_i) \neq y_i} w_{i,t+1}} = \frac{\sum_{h_t(x_i) = y_i} w_{i,t}}{\sum_{h_t(x_i) \neq y_i} w_{i,t}} \text{ at every step, which can simplify the calculation of the new weights.} \]

### 4 Experiments and Results

The proposed framework is collected a human name list. The names in the list are submitted as a query one by one to search for related web images by Google Image Search Engine. The top 200 retrieved web images are crawled automatically. After this OpenCV toolbox is used to detect the faces and adopt the DLK algorithm to align facial images into the same well-define position.

Some baseline annotation methods are used for comparison such as:

- **ORI** – a baseline method that simply adopts the original label information for search-based annotation scheme.
- **CL** – a consistency learning algorithm proposed to enhance weakly labelled image database.
- **MKM** – a modified K-means clustering algorithm to cluster web facial images associated with the extracted names from the surrounding caption.
- **LPSN** – label propagation through sparse neighbourhood algorithm to propagate label information among the neighbourhoods achieved by sparse coding.
- **SLR\(_{\beta=0}\)** – proposed SLR without Sparsity Regularizer.
- **SLR** – proposed supervised label refinement method.

#### 4.1 Evaluation of Facial Feature Representation

The face annotation performance of five types of facial features for the baseline ORI algorithm is evaluated. From the experimental results it is clear to observe that GIST is much or at least slightly better than other common features.

#### 4.2 Evaluation of Auto Face Annotation

The evaluation of auto face annotation performance is based on Search-Based Face Annotation Scheme. It is clear that SLR employs supervised learning to refine labels consistently performs better than ORI using the original weak label. The results validates that proposed SLR can effectively exploit the underlying data distribution of all data examples to refine label matrix and improve the performance of Search-Based Annotation approach.
5 Conclusion and Future Enhancement

The project proposed a search-based face annotation framework, which focused on tackling critical problem of enhancing the label quality and proposed an Supervised Label Refinement Algorithm. To further improve the scalability a clustering-based approximation solution was proposed which successfully accelerated the optimization task without introducing much performance degradation. The future enhancement can be based on a main factor. One the proposed framework does not take into consideration of high resolution images. This can be addressed by employing highly efficient learning techniques which supports high resolution images.

References

G. Krithika, P. Anitha :: Eliminating Duplication in Mining Weakly Labelled Web Facial Images Using Supervised Machine Learning

Biographies

**Mrs.G.Krithika** has received her Diploma in Computer Technology from PSG Polytechnic College in 2003. She received her B.E. degree in Computer Science and Engineering from Sri Ramakrishna Institute of Technology in 2006. She is currently pursuing M.E. Computer Science Engineering from Dr.N.G.P.Institute of Technology. She is currently working as Lecturer in Dr.N.G.P.Institute of Technology, Coimbatore, India. She is the lifetime member of ISTE.

**Mrs.P. Anitha** received her B.E. degree in Computer Science and Engineering from Avinashilingam University in 2006. She also received M.Tech degree from Anna University of Technology, Coimbatore in 2010. She is now pursuing Ph.D in the field of Network Security and Data Mining. She is now working as Assistant Professor in Dr.N.G.P.Institute of Technology, Coimbatore, India. She has published papers in National and International Conferences. She is the lifetime member of ISTE.