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

IJCSMC, Vol. 3, Issue. 5, May 2014, pg.192 – 196

RESEARCH ARTICLE

An Efficient Multilayer Perceptron Neural Network for Mobile Image Annotation on the Cloud

Jai Bharathi[1], Mrs. M. Tamizharasi[2]

PG Student[1], Assistant Professor[2]
Department of Computer Science and Engineering
Dhanalakshmi Srinivasan College of Engineering and Technology
Affiliated to Anna University
East Coast Road, Mamallapuram, Chennai.
jaiitcs@gmail.com

Abstract
With the development of mobile device, large amount of digital images are produced every day. The requirements of effective indexing and searching image are growing rapidly. There is an apparent shift in research from content based image retrieval to automatic image annotation in order to bridge the gap between low level features and high level semantics of image. Automatic Image Annotation techniques facilitate extraction of high level semantic concepts from images by machine learning techniques. So in this research, the association between mobile and cloud opens a new avenue for image annotation, because the heavy computation can be shifted to the cloud for immediately responding user actions. This paper describes and evaluates an automatic image annotation which uses SURF descriptor to select right number of features and right features for annotation. The proposed framework uses Semi-supervised Multilayer Perceptron Neural Network (MLP) in training phase and annotation phase. Finally, the user captures the image from mobile and then forwards it to the cloud and then Annotating the mobile image by the trained MLP. At last, return the annotated image to the mobile.

Keywords
Image Annotation, Support Vector Machine, SURF, Cloud Computing

I. Introduction

TODAY, smart phones equipped with a digital camera have become more and more popular, and personal digital photo is easily produced in massive quantities. Although it is popular to combine time and directory for the photo management in smart phones, it is inconvenient to effectively retrieve photos at the semantic level. Therefore, a large number of image annotation based systems have been developed by utilizing the semantic keywords for the personal photo organization.

Image annotation aims to assign several key words to an image. It is one of the most fundamental research problems in image processing, computer vision and multimedia. Typically, it is accomplished by the following procedures. Given a collection of training images, we first extract visual features to represent these images. Afterward, a set of models are trained based on these images for the subsequent annotation, each of which corresponds to a particular key word (or concept).Given limited computational resource in a mobile, in general, it is impossible to annotate a newly captured image in an online fashion. In addition, the small storage in a mobile does not allow users to maintain a large number of pictures. Therefore, it is essential to consider an alternative way to annotate a mobile image.

Recently, various approaches based on clustering algorithms, Decision tree method, Text based Bayesian method, and Support vector machines (SVM) have been proposed to resolve the image annotation task. Labelled samples are effective to improve the annotation performance. However, it is difficult to get a large number of labeled samples. Thus semi-supervised learning is promising to improve the annotation performance. In this paper, we introduce the Multilayer Perceptron Neural Network for large-scale image annotation.

Cloud computing is rapidly changing the landscape of information technology , as well as attracting more and more academic attentions such as service-oriented science online image processing services handwriting recognition services and storage service for storing scientific data. Since scientific computing usually requires a large number of resources to deliver results for ever-growing problem size in an acceptable time, the cloud computing technology provides significant benefit for scientific computing community by offering cheap alternative to supercomputers, a much more reliable platform than grids, and a much more scalable platform than clusters. In addition, it deals with the trouble raised by client terminals diversity of Mobile Internet comfortably. Thus, it is natural to transmit the mobile image to the cloud for annotation and storage. By utilizing cloud computing, we can easily handle image annotation requests of mobile devices with different operation systems, such as iOS, Android, Windows Phone, and Blackberry. Thus, image annotation can be provided to the end user as a kind of Cloud Software as a service (SaaS). Especially, the proposed Multilayer Perceptron Neural Network which involves large computing resources are ready applicable in the Cloud Infrastructure as a Service (IaaS) platform and can be easily implemented in a parallel fashion to fully utilize the strength of the cloud computing.

II. Annotation System

Due to the various sources of huge quantities of information in the media, the automatic image annotation is an effective technology to improve the image retrieval. The algorithms and systems used for image annotation are commonly divided into these tasks:

- Feature Extraction
- Training Phase in cloud
- Classification and Annotation Phase

The adopted annotation system, in this work, is shown in Fig.1. The system can be divided into two parts: the first one is about a test dataset of annotated images that have been already done by experts (manual annotation). The PASCAL VOC'07 dataset is used that contains social images collected from the Flickr website. This dataset is used for modelling and training the classifier. The second one can be considered as the image annotation, as the topic of this work. To achieve this goal, the query image is firstly compressed using HSC and transmit the compressed image to the cloud then decode the compressed image in cloud segmented into regions that represent objects in the image. Secondly, the feature vectors of each region are computed and extracted from that query image. Those feature vectors are finally feed into the input of a SVM already trained in order to choose the appropriated annotation keywords.

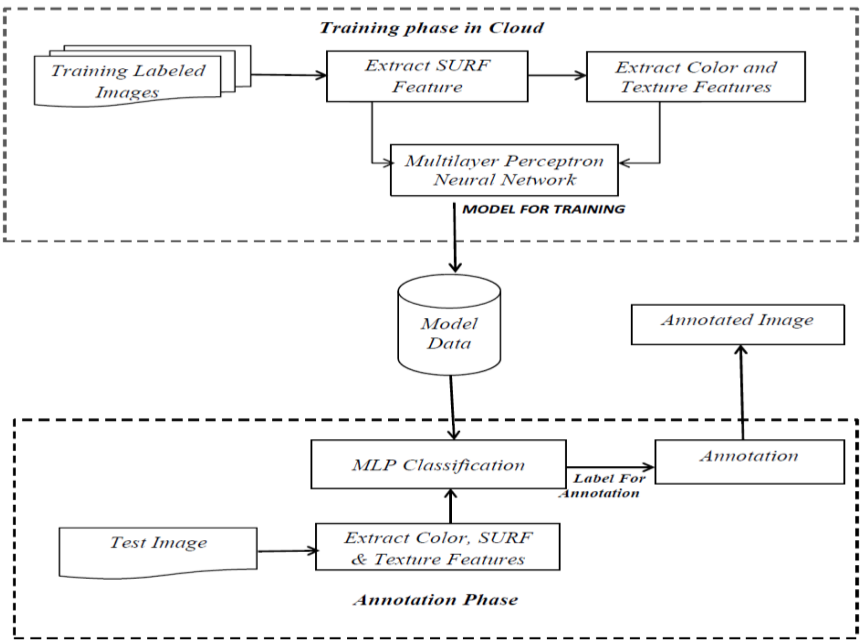

Fig: 1 Annotation System

III. Feature Extraction

A) Colour Descriptor: We want to create a vocabulary to represent the colour of a Training Images. Some images exist in a wide variety of colours, but many have a distinctive colour. Images are often taken in natural outdoor scenes where the lighting varies with the weather and time of day. In addition, images are often more or less transparent, and specular highlights can make the images appear lighter or even white. These environmental factors cause large variations in the measured colour, which in turn leads to confusion between classes. One way to reduce the effect of illumination variations is to use a colour space which is less sensitive to it. Hence, we describe the colour using the HSV colour space. In order to obtain a good generalization, the HSV values for each Pixels in the training images are clustered using k-means clustering. Given a set of cluster centers (visual words) $wc\ i$, $i = 1, 2, ..., Vc$, each image Ij , $j = 1, 2, ...,N$, is then represented by a Vc dimensional normalized frequency histogram $n(wc|Ij)$.

B) Shape Descriptor: The shape of individual image, their configuration, and the overall shape of the images can all be used to distinguish between classes. Changes in viewpoint and occlusions of course change the perceived shape of the image. Thus we need a rotation invariant descriptor. We compute SURF descriptors on a regular grid and optimize over three parameters: the grid spacing M, with a range from 10 to 70 pixels; the support region for the SURF computation with radius R ranging from 10 to 70 pixels; and finally, the number of clusters. We obtain $n(ws|I)$ through vector quantization.

Fig. 2 SURF Features Extraction

C) Texture Descriptors: We describe the texture by convolving the images with MR8 filter bank. The filter bank contains filter at multiple orientation. Rotation invariance is obtained by choosing the maximum response over orientations. We optimize over filters with square support regions of size $s = 3 - 19$ pixels. A vocabulary is created by clustering the descriptors and the frequency histograms $n(wt|I)$ are obtained.

IV. Training Phase in Cloud

The goal of pattern's classification is to allocate an object represented by a number of vector's features into one of a finite set of classes from the reference's database. In order to classify unknown patterns, a certain number of training samples, which are available for each class, are used to train the classifier. The learning task is to compute a classifier or a model that approximate the mapping between the input-output examples and that labels correctly the training set with some levels of accuracy. This can be called the training or model generation stage. After being generated and trained, the model is able to classify an unknown instance, into one of the learned and labeled class in the training set. In other words, the classifier calculates the similarity of all the trained classes and assigns the unlabeled instance to the class with the highest measured similarity.

Consequently, image annotation can be approached by the generated and trained classifier to reduce the gap between low-level vector's feature and high-level concepts. The learned function can directly make the low-level features' set correspond to high-level conceptual classes.

A) Multilayer Perceptron Neural Network:

Multilayer Perceptron Neural networks (or artificial neural networks) are learned by experience. MLP generalize previous experiences to new ones and can make decisions. A neural network can be considered as a black box non-parametric classifier. Neural networks are therefore more flexible. A multilayer network consists of an input layer including a set of input nodes, one or more hidden layers of nodes, and an output layer of nodes. The input layer formed by m nodes, one hidden layer formed by 10 nodes, and output layer formed by n nodes. This neural network is trained to classify inputs according to target classes. The target data should consist of vectors of all zero values except for the one in element i, where i is the represented class. This type of neural network can approximate any function with the acceptable precision. The neural transfer function used, in this tree layer neural network, is a hyperbolic tangent sigmoid transfer function. For image annotation, low-level vector's features are calculated iteratively for each region in the image, either by using SURF visual descriptor. These vector's features are fed into the input layer of the neural network that is already trained, where each of the input neurons or nodes corresponds to each element of these features. And the output neurons of the neural network represent the class labels of images to be classified and annotated. Then each region is annotated by the corresponding label found by neural network classifier.

During the training phase the feature vector extracted from the image dataset are fed into MLP for training the class name of each image is also given during phase. We ran a set of experiments on a private cloud computing platform built on the Amazon system. We used two high performance computer workstations and high speed switch to build the private cloud computing system. The system installed Ubuntu Server.

V. Classification and Annotation Phase

In the annotation phase the user captured image is send to cloud for annotation. In the cloud the MLP is trained using large no.of labelled and unlabelled images. Thus the image annotation is done through MLP in the cloud. First the feature vector is calculated for test image send by the user then it is fed into the MLP as input for classification. The MLP predict the image category and annotation is done and image is again sent to the user.

VI. Dataset

In our image annotation experiments, we use the PASCAL VOC'07 dataset that contains social images collected from the Flickr website. This widely used PASCAL VOC'07 dataset includes 9,963 images which were approximately equally divided into training and test subsets. The images in the dataset were downloaded by querying for images of 20 common entities, including person, bird, cat, cow, dog, horse, sheep, aero plane, bicycle, boat, bus, car, motorbike, train, bottle, chair, dining table, potted plant, sofa, and TV/monitor, from 4 big groups (person, animal, vehicle and indoor). Furthermore, the images were labelled carefully. There are 24,640 annotated objects in the dataset. This indicates that multiple objects from different classes may be presented in the same image. Each image has a corresponding annotation file giving information such as, the relevant filename, source, size, object name and boundary box of object and so on. We mainly use the object name information given by the annotation file during the training process in our experiments.

VII. CONCLUSION

In this paper, we presented an image annotation system using MLP. For this image annotation system, we discussed the effect of using Colour, Shape and Texture descriptors as feature extraction methods. The MLP is chosen and trained to classify and annotate the input image by the suited keywords. The performance of each feature extraction method has been experimentally analysed. The successful experimental results proved that the proposed image annotation system gives good and modest results for some images that are well and properly segmented. Also, the gap between the low-level features and the semantic content of an image must be reduced and be considered for more accuracy of any image annotation system. This system can be also improved by incorporating other classifiers and feature extraction methods can be included and compared to the MLP.

References

[1] *Misra and Kate N. Nagaraj,"Security in Wireless Ad Hoc Networks", in Book The Handbook of Ad-Hoc Wireless Networks (Chapter 30), CRC Press LLC, 2003.*

[2] *M. Armbrust, A. Fox, R. Griffith, A. D. oseph,R.Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, ―A view of cloud computing,‖ Commun. ACM, vol. 53, no. 4, pp. 50–48, 2010.*

[3] *K. Barnard, P. Duygulu, N. de Freitas, D. A. Forsyth, D. M. Blei, and M. I. Jordan, ―Matching words and pictures,‖ J. Mach. Learning Res., vol. 3, pp. 1107–1135, 2003.*

[4] *Dengsheng Zhang, Md. MonirulIslam, GuojunLu, ―A review on automatic image annotation techniques‖, Pattern Recognition, May 2011*

[5] *P. H. Gosselin and M. Cord, ―Active learning methods for interactive image retrieval,‖ IEEE Trans. Image Process., vol. 17, no. 7, pp. 1200–1211, Jul. 2008.*

[6] *Goh K-S, Chang EY, Li B. Using one-class and two-class SVMs for multiclass image annotation. IEEE Trans Know Data Eng 2005;17(10): 1333-1346.*

[7] *J. Sivic, B. Russell, A. Efros, A. Zisserman, and W. Freeman. Discovering object categories in image collections. In Proc. ICCV, 2005.*

Author:

Jai Bharathi.S received her B.E degree in Computer Science and Engineering from Mahendra College Of Engineering, India, in 2012. He is currently pursuing her M.E in Computer science and Engineering from Dhanalakshmi Srinivasan College of Engineering and Technology, India. His research interest includes Image Processing and Cloud Computing.