Chapter 62
Recognition on Images From Internet Street View Based on Hierarchical Features Learning With CNNs

Jian-min Liu
Central South University, China & Hunan Institute of Humanities, Science and Technology, China

Min-hua Yang
Central South University, China

ABSTRACT

This article describes hierarchical features with unsupervised learning on images from internet street view images. This is due to the time spent by trained researchers on feature construction steps with traditional methods. This article focuses on the activation of each layer of with convolutional neural networks (CNNs) on Internet street view images detection and compared similarities and differences among them on each layer. The experiment results achieved error rates of 21% on recognition which work went relatively well than the traditional machine learning techniques, such as Parallel SVM.

INTRODUCTION

Nearer in the last year, artificial intelligence is developing rapidly, which is widely used in areas such as speech recognition, face recognition and web image retrieval.

As everyone knows, and conventional artificial intelligence algorithms follow the following path: pre-process, characteristics construction, characteristics extraction, classification.

A variety of different methods have the same steps as above. The classic methods typically include regression algorithm, k-means, maximum likelihood recognizer, minimum distance recognizer, and SVM.

Like a coin has two sides, everything has its advantages and disadvantages. The general characteristics of these methods are same. On the low level, pixels are extracted from an image and processed, and then the spatial relation of the high-level feature information called edge is further extracted after the completion of pixels processing.

DOI: 10.4018/978-1-5225-8054-6.ch062

Some commonly used edge detection methods include Roberts Cross, Sobel, Canny, Laplacian, Marr-Hildreth, Laplacian of Gaussian, difference of Gaussian. A localized feature extraction also includes harris, change in edge direction, and scale invariant feature transform (SIFT), and so on, and SIFT is one of the most outstanding methods.

The Roberts cross operator (Roberts, 1965) and Prewitt edge detection operator (Prewitt and Mendelsohn, 1966) realize basic first order edge detection. Improved operators include Sobel edge detection operator (Sobel, 1970), Which double the weight of both Prewitt templates, and detect the edge vector.

The Canny edge detection operator (Canny, 1986) is improved based on Gaussian operator, and an advanced implementation (Deriche, 1987) is improved based on Canny.

The most operator is sensitive to changes in image scale and do not work well for matching differing size images. On the one hand, many characteristics are instable when illumination angle, size and scale are changed or image rotation, but on the other hand, some features can not easily distinguish objects from other things.

SIFT (Lowe, 1999, 2004) obtain stable value and separability with high efficiency and accuracy and when illumination angle, size, scale and image rotation as compound independent variable, and next fix many problems in low-level feature extraction. Rather naturally, higher level abstraction benefit from low-level feature extraction.

Nowadays there are all kinds of massive free images via Internet, and program can easily get a large number of pictures with low cost by web spider search. It can be imagined getting training set is very difficult and high cost.

These data sets via Internet street view are very different from manual marked data set, which contains bird, vehicle, flower, animal, road signs, classical architecture, complex street, any other or maybe not. In one word, all of them had not any label.

For example, a classification and recognition tasks, and whose goal is to distinguish between cars and small trucks based on Internet street images. Traditional machine learning methods need a large number of training samples, such as a variety of forms of Internet street images of car which has been manually labeled as a car, and a variety of small Internet street images of truck which has been manually labeled as a small truck.

It is necessary organize a large number of researchers to identify and label for getting features, and getting a large number of training data with accurate tags.

RELATED RESEARCH WORK

Traditional image processing approach work quite well on low resolution remote sensing images(LRRS) but obtain low hit rate high resolution remote sensing images (HRRS). Depend on features template (Zhanga & Zhou, 2004), the researcher achieve image and template matching, and high hit rate of object recognition.

Object recognition of HRRS can depend on intra-domain knowledge provided by experienced experts (Durand, Derivaux, Forestier, Wemmert, Gançarski, Boussaid & Puissant, 2007). The results of experiments show relatively good results and high hit rate when distinguish objects in urban and rural areas.

Different kernel-based approaches of hyperspectral image classification are compared with each other, which include kernel fisher discriminant(KFD), regularized radial basis function neural networks(Reg-RBFNN), regularized AdaBoost (Reg-AB), standard support vector machines (SVMs). The results show